



TEKNILLINEN TIEDEKUNTA

Lämpöpumpun lämpökertoimen ennustaminen datapohjaisella mallinnuksella

Eerik Pyhtilä

PROSESSITEKNIIKAN TUTKINTO-OHJELMA

Diplomityö

Syyskuu 2021



TEKNILLINEN TIEDEKUNTA

Lämpöpumpun lämpökertoimen ennustaminen datapohjaisella mallinnuksella

Eerik Pyhtilä

Ohjaajat: Mika Ruusunen, Jari Ruuska, Petri Hietaharju ja Ari Vuokila

PROSESSITEKNIIKAN TUTKINTO-OHJELMA

Diplomityö

Syyskuu 2021

TIIVISTELMÄ

OPINNÄYTETYÖSTÄ

Oulun yliopisto Teknillinen tiedekunta

| | | | |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------|----------------------------------------------------------------------------------------|-----------------|
| Koulutusohjelma (kandidaatintyö, diplomityö) Prosessitekniikka | | Pääaineopintojen ala (lisensiaatintyö) | |
| Tekijä Eerik Pyhtilä | | Työn ohjaaja yliopistolla Mika Ruusunen, Jari Ruuska, Petri Hietaharju, Ari Vuokila | |
| Työn nimi Lämpöpumpun lämpökertoimen ennustaminen datapohjaisella mallinnuksella | | | |
| Opintosuunta Automaatiotekniikka | Työn laji Diplomityö | Aika Syyskuu 2021 | Sivumäärä 60 |
| Tiivistelmä | | | |
| <p>Tämän työn tavoitteena oli ennustaa lämpöpumpun tulevia tiloja mitatun datan avulla. Lämpöpumput ovat laitteita, joilla hyödynnetään maahan, kallioon, veteen tai ilmaan varastoitunutta lämpöenergiaa rakennusten ja käyttöveden lämmittämiseen.</p> <p>Tarkoituksena oli tutkia, millaisilla datapohjaisilla menetelmillä voidaan ennustaa tehokkaasti lämpöpumppujen toimintaa ja miten voidaan havaita muutoksia niiden tilassa. Yksi keskeinen lämpöpumppujen tehokkuutta kuvaava mittari on lämpökerroin, joka kertoo kuinka paljon enemmän laite tuottaa lämpöä verrattuna sen kuluttamaan sähkömäärään. Työn keskiössä oli lämpökertoimen lyhyen aikavälin ennustaminen datapohjaisilla menetelmillä.</p> <p>Työn teoriaosuudessa käytiin läpi lämpöpumppujen yleistä toimintaa ja datapohjaisia menetelmiä lämpöpumpun lämpökertoimen ennustamiseen liittyen. Useimmat aiemmat tutkimukset liittyivät maa- ja ilmalämpöpumppujen lämpökertoimen ennustamiseen ja yleisimmät käytetyt menetelmät olivat järjestyksessä keinotekoiset neuroverkot, adaptiivinen neurosumea päättelyjärjestelmä, satunnaismetsä-algoritmi, tukivektoregressio ja muut regressiomallit.</p> <p>Kokeellisessa osassa testattiin viittä erilaista dynaamista regressiomallia ja kolmea epälineaarista neuroverkkomallia. Käytössä oli kaksi mittausaineistoa, joista ensimmäinen oli simuloitu ja toinen erään poistoilmalämpöpumpun todellisia mittauksia. Mittausaineistot sisälsivät tulomuuttujina käytetyt ulkolämpötilan, latauspiirin paluuveden lämpötilan ja kompressorin tehon sekä referenssinä käytetyn lämpökertoimen todellisen arvon. Toisella mittausaineistolla jokainen identifioitu malli validoitiin jakamalla mittausaineisto kolmeen osaan, jotka koostuivat 7 vuorokauden tai 14 vuorokauden opetusjaksosta ja niiden jälkeisestä kahden vuorokauden testijaksosta. Mallien ennustetarkkuutta arvioitiin keskineliövirheen neliöjuurella, keskimääräisellä absoluuttisella virheellä ja korrelaatiolla.</p> <p>Kaikki testatut mallit pystyivät toteuttamaan kahden vuorokauden lämpökertoimen ennusteen, mutta niiden tarkkuudessa oli eroja. Parhaat mallit saatiin sekä regressiomalleille että neuroverkoille kaikilla kolmella tulomuuttujalla. Näistä epälineaarinen neuroverkko oli parempi kuin paras regressiomalli. Neuroverkon keskimääräinen absoluuttinen suhteellinen virhe ja korrelaatio olivat testaukseen käytetyllä todellisella mittausaineistolla 6,06 % ja 0,69 ja parhaan regressiomallin vastaavasti 6,13 % ja 0,68. Tulosten perusteella lämpöpumpun toiminnan tehokkuutta pystytään onnistuneesti ennustamaan työssä sovelletuilla datapohjaisilla malleilla.</p> | | | |
| Muita tietoja | | | |

ABSTRACT FOR THESIS

University of Oulu Faculty of Technology

| | | | |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------|--------------------------------------------------------------------------------|-----------------------|
| Degree Programme (Bachelor's Thesis, Master's Thesis) Process engineering | | Major Subject (Licentiate Thesis) | |
| Author Eerik Pyhtilä | | Thesis Supervisor Mika Ruusunen, Jari Ruuska, Petri Hietaharju, Ari Vuokila | |
| Title of Thesis Data-based modelling for predicting efficiency of heat pumps | | | |
| Major Subject Automation Engineering | Type of Thesis Master's thesis | Submission Date September 2021 | Number of Pages 60 |
| <p>Abstract</p> <p>The aim of this work was to predict the future states of the heat pump with the measured data. Heat pumps are devices that utilize thermal energy stored in the ground, rock, water, or air to heat or sometimes also cool buildings and domestic water.</p> <p>The aim was to find answers to question that what kind of data-based methods can be used to effectively predict the operation of heat pumps and how changes in the state of these systems can be detected. One key measure for the efficiency of heat pumps is the coefficient of performance (COP) that is a ratio of useful heating provided to energy required. The core of this work is the short-term prediction of the coefficient of performance using data-based modelling methods.</p> <p>The theoretical part of the thesis reviewed the general theory of heat pumps, the theory of data-based methods and a literature review of previous studies related to the prediction of the COP value. Most of the COP prediction studies were related to ground source heat pumps and air source heat pumps and the most common methods used were in the order artificial neural network, adaptive neuro fuzzy inference system, random forest, support vector regression and other regression model structures.</p> <p>In the empirical part of the thesis, five different dynamic regression models and three nonlinear neural network models were tested. Two datasets were used, the dataset 1 was simulated hourly data, and the dataset 2 was a measured minute level data from the exhaust air heat pump. Both datasets included the outdoor temperature, the supply water temperature and compressor power as input variables, and the measured COP value for the reference. With the dataset 2, each model was validated by dividing the dataset into three parts, each containing 7-day or 14-day training set and after that 2-day test set. The model performances were evaluated with statistical performance indexes: root-mean-square error, mean absolute percentage error, and correlation.</p> <p>All the tested models were able to predict the COP values two-days ahead, but there were differences in their accuracy. The best models were obtained for both regression models and neural networks when all three input variables were selected, and of these, the best neural network model was superior to the best regression model identified. With measured data, the mean absolute percentage error and correlation for test sets were 6.06 % and 0.69 for the neural network and 6.13 % and 0.68 for the best regression model, respectively. According to the results, the state of the heat pump can be successfully predicted with the data-based modelling approach applied in this thesis.</p> | | | |
| Additional Information | | | |

ALKUSANAT

Haluan osoittaa suuret kiitokset työni ohjaajille Mika Ruususelle, Jari Ruuskalle, Ari Vuokilalle ja Petri Hietaharjulle kärsivällisestä ja asianosaavasta ohjeistuksesta työn eri vaiheissa. Erityiskiitoksen ansaitsee etenkin Petri, sillä hänen neuvoista on ollut runsaasti apua työn kokeellisen osan toteutuksessa. Haluan kiittää myös HögforsGST Oy:tä ja Oulun Energia Oy:tä, sillä ne mahdollistivat oikean datan saannin mallinnusta varten.

Tämä diplomityö on toteutettu Oulun yliopiston Ympäristö- ja kemiantekniikan tutkimusyksikössä säätötekniikan ryhmässä osana HOPE-projektia (Highly Optimized Energy Systems).

Oulussa, 9.9.2021

Eerik Pyhtilä

SISÄLLYSLUETTELO

TIIVISTELMÄ

ABSTRACT

ALKUSANAT

SISÄLLYSLUETTELO

MERKINNÄT JA LYHENTEET

| | |
|----------------------------------------------------------------|----|
| 1 JOHDANTO | 9 |
| 2 LÄMPÖPUMPUT | 10 |
| 2.1 Toimintaperiaate | 10 |
| 2.2 Lämpöpumpputyypit ja niiden ominaisuudet..... | 11 |
| 2.2.1 Maalämpöpumppu | 11 |
| 2.2.2 Ilma-ilmalämpöpumppu | 12 |
| 2.2.3 Ilma-vesilämpöpumppu | 12 |
| 2.2.4 Poistoilmalämpöpumppu | 13 |
| 2.3 Lämpöpumput energiajärjestelmän osana..... | 13 |
| 2.4 Tehokkuus ja suorituskkyky | 14 |
| 2.4.1 Lämpökerroin | 14 |
| 2.4.2 Lämpökertoimen ennustamisen hyödyt..... | 15 |
| 3 DATAPOHJAINEN MALLINNUS | 16 |
| 3.1 Mallin identifiointi | 16 |
| 3.1.1 Datan esikäsittely | 17 |
| 3.1.2 Muuttujien valinta..... | 17 |
| 3.1.3 Mallirakenne | 18 |
| 3.1.4 Opetus ja testaus | 20 |
| 3.1.5 Suorituskyvyn arviointi | 20 |
| 3.2 Datapohjaisten mallinnusmenetelmien rajoitteita | 21 |
| 4 LÄMPÖKERTOIMEN ENNUSTAMISMENETELMIÄ | 23 |
| 4.1 Keinotekoiset neuroverkot | 23 |
| 4.1.1 Verkon rakenne ja toiminta | 23 |
| 4.1.2 Adaptiivinen neurosumea päättelyjärjestelmä..... | 25 |
| 4.1.3 Radiaalinen kantafunktioverkko | 25 |
| 4.1.4 Neuroverkkomalleihin perustuvat aiemmat tutkimukset..... | 26 |
| 4.2 Satunnaismetsä..... | 29 |
| 4.3 Tukivektoregressio | 30 |

| | |
|------------------------------------------------------------------|----|
| 4.4 Muita menetelmiä..... | 31 |
| 4.5 Yhteenveto menetelmistä lämpökertoimen ennustamisessa | 34 |
| 5 LÄMPÖKERTOIMEN MALLINNUS JA SIMULOINTI..... | 37 |
| 5.1 Mittausaineistot..... | 37 |
| 5.2 Lämpökertoimen ennustamiseen sovelletut mallirakenteet | 39 |
| 5.2.1 Regressiomallit | 39 |
| 5.2.2 Epälineaariset autoregressiiviset neuroverkkomallit | 41 |
| 5.3 Tulokset ja niiden tarkastelu | 42 |
| 5.3.1 Regressiomallit: simuloitu data | 42 |
| 5.3.2 Regressio- ja neuroverkkomallit: mittausaineisto | 43 |
| 5.4 Mallinnusvirheen analyysi | 48 |
| 6 JOHTOPÄÄTÖKSET JA SUOSITUKSET | 51 |
| 7 YHTEENVETO | 54 |
| LÄHDELUETTELO..... | 55 |

MERKINNÄT JA LYHENTEET

| | |
|----------------|------------------------------------------------------------------------|
| ANFIS | Adaptiivinen neurosumea päättelyjärjestelmä |
| ANN | Keinotekoinen neuroverkko (Artificial Neural Network) |
| APSO | Adaptiivinen partikkeliparviontimointi |
| AR | Autoregressiivinen malli |
| AV | Absoluuttinen virhe |
| BPNN | Vastavirtaneuroverkko (Backpropagation Neural Network) |
| COP | Lämpökerroin (Coefficient of Performance) |
| COV | Kovarianssi |
| CVRMSE | Keskineliövirheen neliöjuuren vaihtelukerroin |
| FFNN | Eteenpäin syöttävä neuroverkko (Feed Forward Neural Network) |
| FWP | Sumeapainotettu esikäsittely (Fuzzy weighted pre-processing) |
| MAE | Keskimääräinen absoluuttinen virhe |
| MAPE | Keskimääräinen absoluuttinen suhteellinen virhe |
| MBE | Keskivirhe |
| MLFFN | Monikerroksinen eteenpäin syöttävä verkko |
| MLP | Monikerroksinen perseptroni (Multilayer Perceptron) |
| MLR | Lineaarinen moniregressio |
| MRE | Keskimääräinen suhteellinen virhe |
| MRMR | Vähimmäisredundanssin maksimaalinen relevanssi |
| MSE | Keskimääräinen neliövirhe |
| NARX | Epälineaarinen autoregressiivinen eksogeeninen neuroverkkomalli |
| NMBE | Normalisoitu keskivirhe |
| PCA | Pääkomponenttianalyysi |
| r | Korrelaatio |
| RBFNN | Radiaalinen perustoimintoverkko (Radial Basis Function Neural Network) |
| RF | Satunnaismetsä (Random Forest) |
| RMSE | Keskineliövirheen neliöjuuri |
| R ² | Selitysaste |
| SV | Suhteellinen virhe |
| SVM | Tukivektorikone (Support Vector Machine) |
| SVR | Tukivektoriregressio (Support Vector Regression) |

| | | | |
|-------------|---------------------------------------------------------------|----------------|--|
| b_l | Viive | | |
| $c_{p,l}$ | Lämmönsiirtoaineen ominaislämpökapasiteetti | lauhduttimessa | |
| | [kWh/kgK] | | |
| f | Aktivaatiofunktio | | |
| P | Kompressorin hetkellinen teho [kW] | | |
| P_{el} | Lämpöpumpun kuluttama sähköteho [kW] | | |
| T_B | Latauspiirin paluuvien lämpötila [°C] | | |
| $t_{h,in}$ | Tulolämpötila höyrystimeen [K] | | |
| $t_{l,out}$ | Lauhduttimen ulostulon lämpötila [K] | | |
| T_{out} | Ulkolämpötila [°C] | | |
| \dot{V} | Tilavuusvirtaus lauhduttimessa [m ³ /h] | | |
| w_{il} | Painokerroin | | |
| x_i | Tulosignaali | | |
| ρ_l | Lämmönsiirtoaineen tiheys lauhduttimessa [kg/m ³] | | |

1 JOHDANTO

Euroopan Unionin vuoteen 2030 tähtäävä energia- ja ilmastopaketti velvoittaa myös Suomea energiasysteemien tehostamiseen ja ilmastopäästöjen vähentämiseen (Maa- ja metsätalousministeriö 2021). Monimutkaistuvan energiasysteemin ja sen osien hallinta vaatii uusia työkaluja, joilla tehokkuutta voidaan nostaa.

Yksi keino on löytää entistä parempia keinoja energialaitteiden ja prosessien ennustamiseen ja säätöön hyödyntämällä dataa eri lähteistä, kuten laitteesta, kuluttajilta ja säätiedoista. Kehittyneessä energiajärjestelmässä data on tärkeässä roolissa tehokkaan operoinnin kannalta. Mitä paremmin laitteiden toimintaa voidaan ennustaa, sitä paremmin voidaan oppia hallitsemaan niiden optimointiin liittyviä epävarmuuksia ja sitä paremmin opitaan säätämään niiden toimintaa energiaa säästävään suuntaan.

Lämpöpumput ovat lämmityksen toteutuksessa kasvattaneet runsaasti suosiota, ja tulevaisuudessa niiden merkitys voi kasvaa entisestään. Lämpöpumput ovat laitteita, joilla voidaan tuottaa rakennusten tarvitsemaa lämpöä ympäristöstä saatavasta lämpöenergiasta. Lämpö voidaan ottaa esimerkiksi maasta, vedestä tai ilmasta.

Tämän työn tavoitteena on ennustaa lämpöpumpun toimintaa mittausten avulla. Tarkoituksena on tutkia, millaisilla datapohjaisilla menetelmillä voidaan ennustaa tehokkaasti lämpöpumppujen toimintaa ja miten voidaan havaita muutoksia niiden tilassa. Keskeinen lämpöpumppujen tehokkuutta kuvaava mittari on lämpökerroin, joka kertoo kuinka paljon enemmän laite tuottaa lämpöä verrattuna sen kuluttaman sähkön määrään. Lämpökertoimen lyhyen aikavälin ennustaminen datapohjaisilla menetelmillä on siten tämän työn keskiössä. Työssä testatut mallit on toteutettu poistoilmalämpöpumpun ja ilma-vesilämpöpumpun datoilla. Työssä käytettiin mallirakenteina dynaamisia regressiomalleja ja epälineaarisia autoregressiivisiä eksogeenisiä neuroverkkomalleja.

2 LÄMPÖPUMPUT

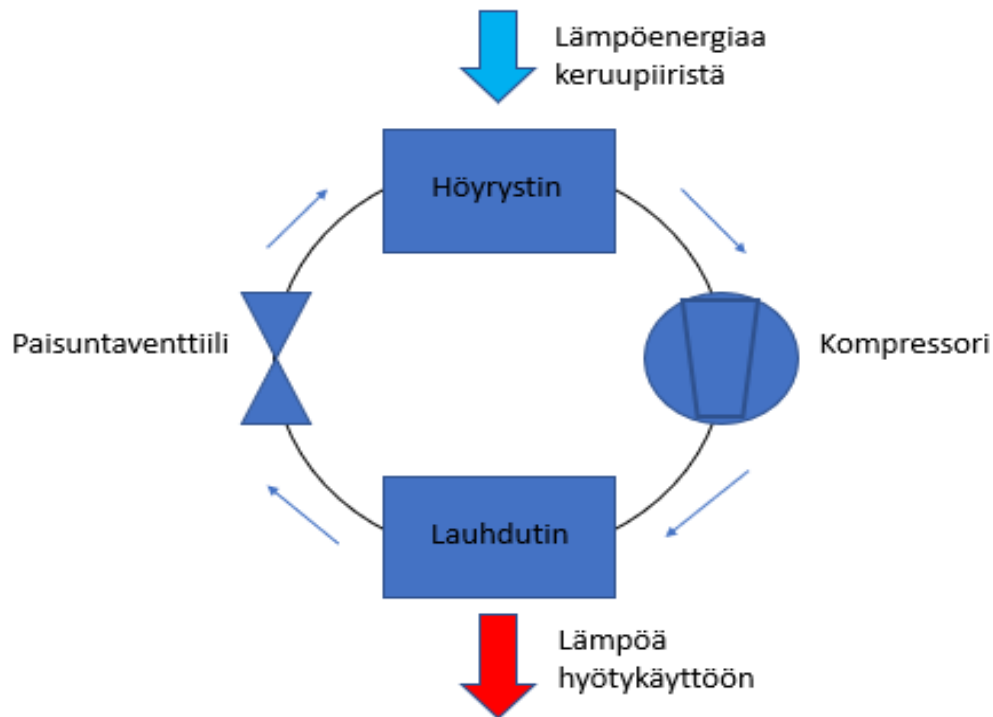
Lämpöpumput ovat laitteita, joiden tekemän työn avulla hyödynnetään maahan, kallioon, veteen tai ilmaan varastoitunutta lämpöenergiaa rakennusten ja käyttöveden lämmittämiseen. Lämpöpumppuja on pidetty pitkään tulevaisuuden puhtaan energian lähteenä rakennusten lämmöntarpeeseen, jos niiden yhteydessä voidaan hyödyntää uusiutuvaa energiaa (Reynolds ym. 2018).

Maalämpöpumput hyödyntävät maaperästä, kalliosta tai vedestä saatavaa lämpöenergiaa. Ilmalämpöpumput puolestaan käyttävät ilmasta saatavaa lämpöenergiaa. Maalämpöpumppuja ja ilmalämpöpumppuja voidaan hyödyntää myös viilennystarkoitukseen. (Suomen lämpöpumppuyhdistys 2021)

Ilmalämpöpumput hyödyntävät ilmasta saatavaa lämpöenergiaa lämmityskohteessa rakennuksen sisäilman tai vesikiertojärjestelmän lämmitykseen. Ilmalämpöpumppuja ovat ilma-ilmalämpöpumppu, ilma-vesilämpöpumppu ja poistoilmalämpöpumppu. (Motiva 2012b)

2.1 Toimintaperiaate

Lämpöpumppujen toiminta perustuu suljettuun kiertoon, jossa kylmäaine muuttaa olomuotoaan kaasusta nesteeksi ja toisin päin. Lämpöpumpun keskeisimmät osat ovat kompressor, paisuntaventtiili ja kaksi lämmönvaihdinta eli höyrystin ja lauhdutin (Kuva 1). Lämpöenergiaa otetaan talteen höyrystimenä toimivan lämmönvaihtimen avulla lämpöpumpun kylmäainekierto. Kylmäaineen mukana lämpö siirtyy kompressorille, joka puristaa kylmäainehöyryä suuressa paineessa, jolloin höyry tiivistyy takaisin nesteeksi. Syntynyt lämpöenergia hyödynnetään lauhduttimen kautta. Prosessi alkaa alusta, kun nestemäinen kylmäaine kulkee painetta ja lämpötilaa laskevan paisuntaventtiilin kautta jälleen höyrystimelle. (Motiva 2018)



Kuva 1. Lämpöpumpun yksinkertaistettu toimintaperiaate. Nuolien suunta osoittaa kylmäaineen kiertosuunnan. (Mukaiillen Motiva 2012a)

2.2 Lämpöpumpputyypit ja niiden ominaisuudet

2.2.1 Maalämpöpumppu

Maalämpöpumput hyödyntävät maahan, kallioon tai veteen varastoitunutta auringon lämpöä. Maalämpöpumpun tuottama energia on pääosin uusiutuvaa energiaa, joten maalämpöpumput ovat varteenotettavia ympäristöystävällisiä vaihtoehtoja lämmön tuottamiseen. Lämpö otetaan joko pintamaahan asennetusta pitkästä vaakaputkistosta, lämpökaivosta eli syvästä porakaivosta tai vedestä. Maaperästä tai kalliosta kerätyn energian lämpötilaa nostetaan kompressorin avulla korkeaksi jopa sataan asteeseen. Lämpökaivon syvemmissä osissa lämpöä saadaan myös maapallon ytimeistä sekä lämpimistä pohjavesivirtauksista. Yleisin maalämpöä hyödyntävä kohde toteutetaan lämpökaivoilla, jossa porakaivoon asennetaan putkisto, jossa lämmönkeruuliuos kiertää. Maalämpöpumpun lämpökerroin on sitä parempi, mitä pienempi lämpötilaero lämmönlähteen, kuten maaperän ja lämpöä kohteeseen luovuttavan patterin tai putkiston välillä on. (Motiva 2012a; 2018)

Lämpökaivon lämmönkeruuputkistossa kiertää lämmönkeruuneste, joka on useimmiten 70 % vettä ja 30 % bioetanolia. Maalämpöä saadaan kerättyä energiakaivosta, kun maaperä lämmittää lämmönkeruunestettä. (Thermia 2021)

Täysteholle mitoitettu maalämpöpumppu pystyy kattamaan koko rakennuksen lämmitystarpeen myös kovimmilla pakkasilla. Lämmitysenergian tuottaminen on sillä edullista, mutta maalämpöpumpun alkuinvestointi on kohtalaisen kallis, joten maalämpöpumppuja käytetään useimmiten suuremmissa rakennuksissa. Pumppujen teknologian kehittyminen sekä sähkön ja öljyn hinnan nousu ovat vaikuttaneet maalämmön yleistymiseen myönteisesti. (Motiva 2018)

2.2.2 Ilma-ilmalämpöpumppu

Ilma-ilmalämpöpumppu on tavallisesti lisälämmityslaite, jota voidaan käyttää päälämmitysjärjestelmän tukena isommissa rakennuksissa. Pienemmissä kohteissa se voi toimia myös päälämmittäjänä. Ne lämmittävät sisätilojen ilmaa siirtämällä lämpöä ulkoilmasta. Ilma-ilmalämpöpumppua ei käytetä käyttöveden lämmittämiseen eikä sitä voida liittää vesikiertoiseen lämmitysjärjestelmään, kuten ilma-vesilämpöpumppua. Ilma-ilmalämpöpumput soveltuvat myös rakennuksen viilentämiseen. (Motiva 2012b)

Ilma-ilmalämpöpumpuissa lämpö otetaan ulkoilmasta, jolloin lämpöpumpun suorituskyky riippuu huomattavasti ulkolämpötilasta. Ilma-vesilämpöpumpun tehokkuuteen vaikuttaa rakennuksen eristysten taso ja käytetty lämmönjakojärjestelmä. Myös rakennuksen pohjarakenteen avaruudella on merkitystä. (Motiva 2012b)

Kova pakkas on vaikuttaa ilmalämpöpumpun lämpökertoimeen heikentävästi. VTT:n testiraportin (2018) mukaan ilmalämpöpumpun lämpökerroin oli kovimmilla pakkasilla 1,5–2, kun taas parhaimmillaan se oli noin 6–7,5, ulkolämpötilassa 10 °C.

2.2.3 Ilma-vesilämpöpumppu

Ilma-vesilämpöpumppujen tekemällä työllä lämpöä siirretään ulkoilmasta vesikiertoiseen lämmitysjärjestelmään ja niitä käytetään myös käyttöveden lämmittämiseen. Järjestelmä voidaan liittää joko vesikiertoiseen lattialämmitykseen tai pattereihin. (Motiva 2012b)

Ilma-vesilämpöpumpputjärjestelmään kuuluu ulko- ja sisäyksikkö. Ulkoyksikössä on lämpöä keräävä höyrystin ja kompressor. Joissakin malleissa (monoblock) lämpöä luovuttava lauhdutin sisältyy myös ulkoyksikköön. Sisäyksikköön kuuluu ohjainyksikkö ja usein myös lauhdutin. (Motiva 2018)

2.2.4 Poistoilmalämpöpumppu

Poistoilmalämpöpumppu hyödyntää rakennuksen ilmanvaihdon kautta poistuvaa lämmintä sisäilmaa. Poistoilmalämpöpumpun tuottama lämpö johdetaan rakennuksen lämmitettäviin tiloihin, vesikiertoon ja käyttöveteen. Niiden teho on tavallisesti pienempi kuin maa- ja ilmalämpöpumpuilla. Jos poistoilman lämpö ei riitä kattamaan koko rakennuksen lämmitystarvetta, niin lisälämpöä voidaan tarvittaessa tuottaa sähkövastuksilla. Poistoilmalämpöpumppu on muihin ilmalämpöpumppeihin verrattuna riippumattomampi vallitsevasta säästä, sillä lämpö otetaan poistuvasta sisäilmasta. Järjestelmään kuuluu tavallisesti puhaltimen tai puhaltimia sisältävä lämmöntalteenottolaite, lämmönkeruuputkisto, lämpöpumppu, vesivaraaja ja lisälämmönlähde. (Motiva 2012b)

2.3 Lämpöpumput energiajärjestelmän osana

Lämpöpumppujen merkityksen oletetaan kasvavan entisestään tulevaisuudessa kysyntäjoustotarpeen kasvaessa. Tulevaisuudessa hyödynnetään yhä enemmän sääolosuhteista riippuvaista uusiutuvaa sähköenergiaa, kuten aurinko- ja tuulivoimaa. Sähkömarkkinoiden tulee kyetä sopeutumaan vaihtuvista sääolosuhteista johtuviin vaihteluihin energiantuotannossa. Lämpöpumppuja voidaan käyttää apuna energian varastoinnissa kysynnän ollessa vähäistä ja puolestaan tuottaa niillä energiaa lämpöverkkoon energian hinnan ollessa korkealla eli kysynnän ollessa suuri. (Motiva 2020)

Lämpöpumput ovat usein yhdistetty lämminvesivaraajaan, jonka tarkoituksena on parantaa suorituskykyä ja tasata lämpökuormaa. Lämpökerroin on korkea tyypillisesti kuormituspohjan aikana. Silloin lämpöä kannattaa varastoida ja hyödyntää myöhemmin huippukuormituksen aikaan sähkönkulutuksen minimoimiseksi. (Rastegarpour ym. 2020)

2.4 Tehokkuus ja suorituskyky

2.4.1 Lämpökerroin

Lämpöpumpun tehokkuutta kuvataan lämpökertoimella (engl. coefficient of performance, COP). Lämpökerroin saadaan lämpötehon ja sähkötehon suhteesta. Se kertoo kuinka paljon enemmän laite tuottaa lämpöenergiaa verrattuna sen kuluttaman sähkömäärään. Esimerkiksi jos lämpökerroin on kolme, lämpöpumppu tuottaa lämpöä kolme yksikköä jokaista siirtotyöhön käytettyä sähkötehoyksikköä kohden. Paras lämpökerroin saavutetaan, kun lämpötilaero on mahdollisimman pieni lämmönkeruun ja -luovutuksen välillä. (Motiva 2020)

Lämpökerroin voidaan laskea seuraavalla kaavalla (Jesper ym. 2021):

$$COP = \frac{\dot{Q}_l}{P_{el}} = \frac{(t_{l,out} - t_{h,in}) \cdot \dot{V}_l \cdot \rho_l \cdot c_{p,l}}{P_{el}}, \quad (1)$$

missä \dot{Q}_l on hyödynnettävä lämpö [kW],
 P_{el} on lämpöpumpun kuluttama sähkö [kW],
 $t_{l,out}$ on lauhduttimen ulostulon lämpötila [K],
 $t_{h,in}$ on tulolämpötila höyrystimeen [K],
 \dot{V} on tilavuusvirtaus lauhduttimessa [m³/h],
 ρ_l on lämmönsiirtoaineen tiheys lauhduttimessa [kg/m³] ja
 $c_{p,l}$ on lämmönsiirtoaineen ominaislämpökapasiteetti lauhduttimessa [kWh/kgK].

Lämpöpumpun lämpökerroin voi vaihdella vuoden aikana paljon riippuen kohteesta ja olosuhteista. Lämpökertoimille on omat tietyt olosuhteet määrittelevät standardit, joiden mukaan se lasketaan. Yleisen standardin perusteella laskettu lämpökerroin ei kuitenkaan ota huomioon käyttöveden tuotantoa, jossa lämpökerroin on aina pienempi kuin lämmöntuotanto lämmitysverkkoon. Eri lämpöpumppumalleilla lämpökertoimet eivät ole välttämättä vertailukelpoisia, jos on käytetty eri standardia. Lämpöpumppujen mittausstandardit antavat usein liian optimistisen kuvan Suomen ilmasto-olosuhteissa. (Motiva 2018)

Lämpöpumppujen COP-arvot tuotetaan tavallisesti virallisen SFN-EN 14511 standardin mukaisissa testausolosuhteissa. Siinä määritellään ulkoilmalämpöpumppujen osalta ulkoilman lämpötilaksi +7 °C ja sisäilman lämpötilaksi +20 °C. Maalämpöpumpuille puolestaan on määritelty keruupiirin lämpötilaksi 0 °C ja lämmönjakoverkoston menoveden lämpötilaksi +35 °C. (Ympäristöministeriö 2012)

2.4.2 Lämpökertoimen ennustamisen hyödyt

Lämpöpumppusysteemeissä COP on tärkeä tunnusluku tehokkaalle operoinnille. Siihen kuitenkin vaikuttaa useita eri muuttujia ja niistä voi olla vaikeaa saada reaaliaikaisia tietoja, joka olisi tärkeää tehokkaamman energiaa säästävän operoinnin ja optimoinnin kannalta. Ongelmana onkin usein mittauslaitteiston tai monitoroinnin puute. Lämpöpumpun ja koko lämpöpumppusysteemin lämpökertoimen laskeminen perinteisillä menetelmillä voi olla työlästä ja hankalaa, jos tarvittavaa dataa ei ole olemassa. Lämpökertoimen ennustamisessa tulisikin käyttää sellaisia muuttujia, joita on helppo valvoa. (Shin ym. 2019)

Lämpökerroin ei ole vakio, vaikka se sellaisena usein yksinkertaisuuden vuoksi kuvataan. Päinvastoin lämpökertoimen arvo muuttuu dynaamisesti ja siihen vaikuttaa useita tekijöitä, jotka tarkka ennustusmalli pystyy kuvaamaan. Maalämpöpumpun lämpökertoimeen vaikuttavat esimerkiksi kuormitusprosentit, ulkoilman hetkellinen lämpötila ja maapohjan lämpötila. Haasteena onkin saada yhä realistisempia malleja, että entistä tarkempi lämpöpumppujen toiminnan ennustaminen ja energiankulutuksen optimointi olisi mahdollista monitavoitteisessa ympäristössä, jossa energijärjestelmän kysyntä ja tarjonta kohtaavat. (Reynolds ym. 2018)

Lämpöpumppujen lämpökertoimeen vaikuttaa myös lämmönlähteen lämpötila, joka riippuu pitkälti sen ympäristön lämpötilasta. Lämpökerroin riippuu myös lämmönjakopuolen lämpötilasta (veden lämpötila lauhduttimessa) ja kompressorin tehosta. Etenkin ilmalämpöpumppujen suorituskykyyn vaikuttaa myös ulkoilman kosteus, jonka takia jäätä voi muodostua laitteiston höyrystimeen. (Rastegarpour ym. 2020)

3 DATAPOHJAINEN MALLINNUS

Matemaattiset mallit jaetaan usein fysiikkaan pohjautuviin white-box -malleihin, datapohjaisiin black-box -malleihin ja näitä yhdistäviin gray-box -malleihin. Tässä työssä päähuomio on datapohjaisissa malleissa eli black-box -tyyppisissä menetelmissä.

Datapohjaiset mallit perustuvat tilastollisiin ja tekoälytyökaluja hyödyntäviin menetelmiin. Useimmat mallinnustekniikat perustuvat aikasarjojen data-analyysiin ja joissakin menetelmissä tarvitaan lisäksi perustietämystä mallinnuksen kohteesta. (Bourdeau ym. 2019; Bishop 2006)

Black-box -menetelmät ovat mallintamiseen soveltuvia laskennallisia menetelmiä, joissa käytetään systeemistä mitattua historiallista dataa ja sovitetaan siihen matemaattisia funktioita. Siinä ei vaadita samanlaista ymmärrystä systeemin fysiikasta, kuten white-box -malleissa. Black-box -mallien etuna pidetään niiden tarkkuutta verrattuna white-box -malleihin. Niiden haasteena puolestaan pidetään sitä, että mallit eivät ole aina helposti yleistettävissä. (Afram & Janabi-Sharifi 2015)

Datapohjaisten menetelmien tavoitteena on tässä tapauksessa rakentaa malli, joka kuvaa vuorovaikutuksia tulomuuttujien ja halutun lähdön tai lähtöjen välillä jossakin tietyssä prosessissa. Lähtö on se, mitä halutaan ennustaa ja tuloina käytetään datasta eroteltuja muuttujia ja niiden piirteitä. Tulomuuttujien ja piirteiden valintaan datajoukosta on erilaisia strategioita. (Bourdeau ym. 2019)

3.1 Mallin identifiointi

Datapohjaisen mallin identifiointiin kuuluu tyypillisesti seuraavia, osin toistuvia vaiheita (Mathworks 2021):

1. Datan esikäsittely,
2. Muuttujien tai piirteiden muodostaminen ja valitseminen,
3. Mallin rakenteen valinta,
4. Mallin opetus ja testaus ja
5. Mallin suorituskyvyn arviointi.

3.1.1 Datan esikäsittely

Datan esikäsittely on tärkeä vaihe mallin rakentamisessa, koska se voi vaikuttaa huomattavasti ennusteen tarkkuuteen ja luotettavuuteen. Tämän vaiheen tarkoituksena on parantaa datan laatua ja se tähtää optimaaliseen tulomuuttujien, aikaikkunoiden ja aika-askeleiden valintaan (Bourdeau ym. 2019). Vaiheeseen kuuluu datan siistiminen, dataintegrointi, datan muuttaminen ja/tai datan redusointi. Joskus siihen voi kuulua myös datan esivalinta. Datan siistimisellä tarkoitetaan epätäydellisen, väärän, epätarkan tai muuten epäsovikivan datan havaitsemista ja korjaamista esimerkiksi muokkaamalla, täydentämällä, korvaamalla tai poistamalla mittauspisteitä. Dataintegroinnilla tarkoitetaan eri lähteistä tulevien datojen yhdistämistä. Datan muuttaminen on prosessi, jossa data muutetaan oppimisalgoritmillemme sopivampaan muotoon. Se voi sisältää esimerkiksi normalisointia, datan yhdistämistä tai hajottamista sekä datan yleistämistä. Datan redusoinnilla taas voidaan pienentää datajoukon dimensiota, joka voi tehostaa mallin suorituskykyä. Datan redusointiin voidaan käyttää esimerkiksi pääkomponenttianalyysia (engl. principal component analysis, PCA), jolla on samalla mahdollista löytää datasta merkittävimmät piirteet. (Amasyali & El-Gohary 2018)

3.1.2 Muuttujien valinta

Datajoukko voi sisältää merkityksettömiä muuttujia, ylimääräistä kohinaa tai ylimääräistä toistoa. Mallin parhaan suorituskyvyn ja tehokkuuden kannalta datasta täytyy erottaa epäolennaiset piirteet pois ja valita mukaan vain tärkeitä muuttujia. Piirteiden valinta on datankäsittelyyn liittyvä prosessi, jossa kohinaisesta datajoukosta valitaan tai muodostetaan sopivia piirteitä automaattisesti tai manuaalisesti sen mukaan, mitkä piirteet tuottavat parhaan ennustuksen. Se voi myös helpottaa malliin yleistettävyyttä ja vähentää ylisovituksen riskiä. (ScienceDirect, 2021)

Sopivien tulomuuttujien valinta on tärkeä vaihe mallin rakentamisessa. Muuttujien valintavaihe vähentää myös laskenta-aikaa, laskentakapasiteetin tarvetta, parantaa mallin tarkkuutta, tekee siitä yksinkertaisemman sekä nopeuttaa tulosten saamista (Yan ym. 2016).

Jos tulomuuttujien määrä on liian pieni, seurauksena voi olla alisovitusongelma. Toisaalta jos muuttujia on liikaa, mallista tulee monimutkainen ja mahdollisesti ylisovittunut. Kun valitaan sopiva määrä oikeanlaisia muuttujia tai piirteitä, mallinnettavaa systeemiä

pystytään ymmärtämään paremmin ja on myös helpompaa selvittää matemaattisesti, onko valittu piirre merkityksellinen mallinnustuloksen kannalta. (Wang ym. 2019)

Piirteiden valinta voi tähdätä myös suuren datajoukon dimensioiden vähentämiseen. Siihen on olemassa useita eri algoritmeja. Näiden algoritmien tehtävänä on etsiä datajoukosta sellainen osajoukko, jolla voidaan mallintaa optimaalisesti mitatun lähdön rajoitukset huomioiden. Piirteiden valintaan käytettävät algoritmit voidaan jaotella luokitteluun perustuviin (engl. filters), käärealgoritmeihin (engl. wrappers) ja upotettuihin algoritmeihin (engl. embedded). (MathWorks 2020)

Luokitteluun perustuvissa menetelmissä piirteiden tärkeys perustuu niistä laskettaviin ominaisuuksiin, kuten varianssiin sekä piirteen riippuvuudesta lähtöön. Tässä tapauksessa piirteet valitaan datan esikäsittelyvaiheessa ja määritetään malli valituilla muuttujilla, joten valinta suoritetaan opetusalgoritmista erillään. Wrapper-menetelmässä taas piirteiden valinta perustuu valintakriteeriin, joka mittaa muutoksia esivalitun mallin suorituskyvyssä. Algoritmi toistaa opettamista ja parantaa mallia lisäämällä tai poistamalla muuttujia tai piirteitä, kunnes kriteeri saavutetaan. (MathWorks 2020)

Upotusmenetelmässä piirteiden tärkeys opitaan samalla, kun mallia opetetaan. Algoritmi valitsee piirteitä, jotka sopivat hyvin tiettyyn oppimisprosessiin. Piirteiden valintaan käytettävät algoritmit voidaan luokitella myös sen mukaan, että luokittelevatko ne piirteitä peräkkäin vai ei. Peräkkäisten piirteiden valintaan käytettäviä algoritmeja ovat esimerkiksi MRMR-algoritmi eli vähimmäisredundanssin maksimaalinen relevanssi ja vaiheittainen regressio. (MathWorks 2020)

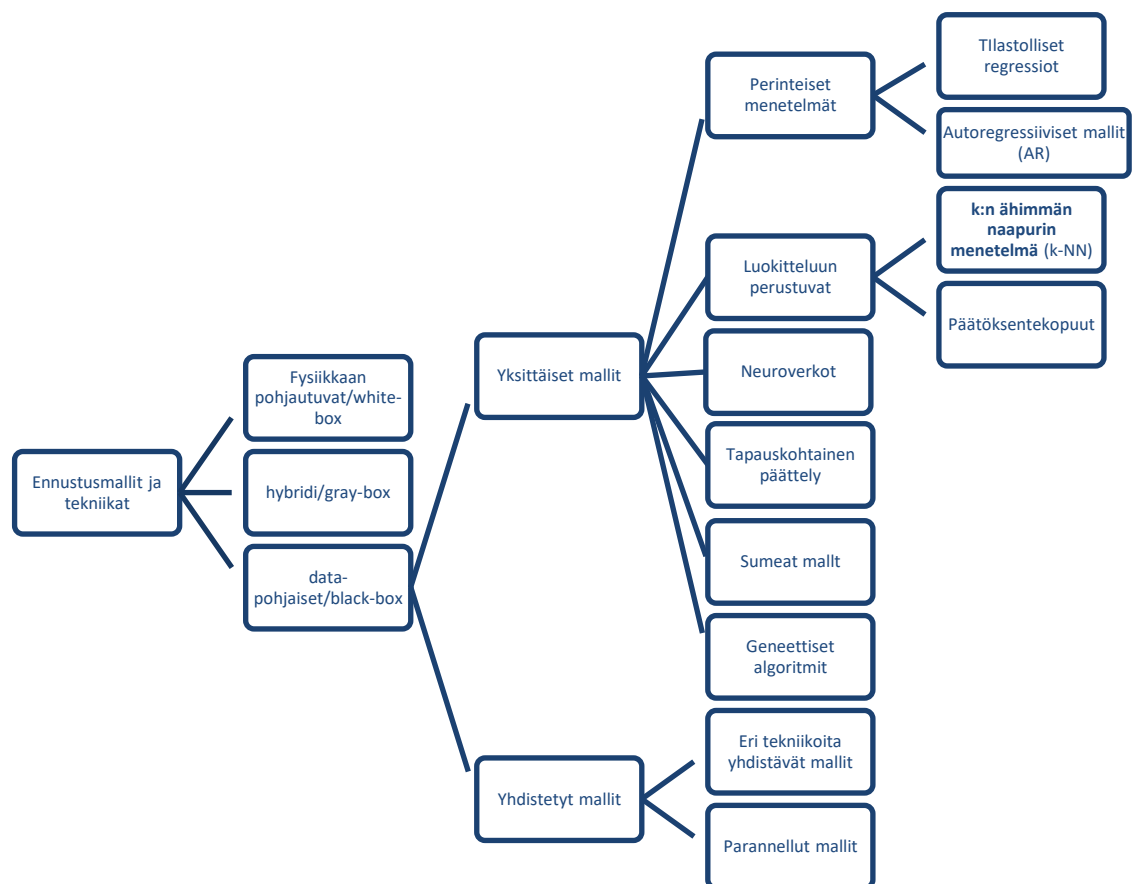
3.1.3 Mallirakenne

Mallin valinnassa on tavoitteena valita paras malli erilaisten mallirakenteiden joukosta. Erilaisia oppimisalgoritmeja on lukuisia, eikä parhaan mallin valinta ole aina helppoa. Valintaan vaikuttavat lukuisat tekijät, kuten ongelman luonne, haluttu vaste, datan tyyppi, laskenta-aika, piirteiden lukumäärä ja datasta tehdyt havainnot. (Bourdeau ym. 2019)

Yksi merkittävä mallirakenteen valintaan vaikuttava tekijä on tasapaino sopivan tarkkuuden ja tulkittavuuden välillä. Malli voi olla samaan aikaan tarkka, mutta kompleksinen ja vaikeasti tulkittava (Srinivasan 2019).

Jokaisella mallilla on omat vahvuutensa ja heikkoutensa. Esimerkiksi jotkut mallirakenteet sopivat paremmin pienelle datajoukolle ja toiset taas suurempiin datajoukkoihin. Eri mallien ennustustarkkuus voi tästä syystä vaihdella. Erilaisten mallien hyvyttä ja sopivuutta tarkasteltavan ongelman ratkaisuun voidaan selvittää mallin validoinnilla, jossa mallin suorituskyky varmistetaan ja se kuvaa tarkasti mallinnettavaa kohdetta. (Bourdeau ym. 2019)

Tyypillisiä datapohjaisia mallirakenteita ovat perinteiset puhtaasti tilastollisiin menetelmiin perustuvat regressiomenetelmät ja erilaiset autoregressiiviset mallit, keinotekoiset neuroverkot (engl. artificial neural networks, ANN), tukivektorikoneet (engl. support vector machine, SVM) ja satunnaismetsä (engl. random forest, RF) algoritmi (Reynolds ym. 2018). Näistä SVM, ANN, päätöspuut ja muut tilastolliset algoritmit ovat yleisimmin käytettyjä ohjattuja koneoppimisalgoritmeja mallinnuksessa (Amasyali & El-Gohary 2018). Bourdeau ym. (2019) käyttivät eri mallirakenteista kuvan 2 mukaista luokittelutapaa rakennusten energiasysteemien ennustamisessa.



Kuva 2. Mallinnusmenetelmien eräs luokittelutapa (mukaillen Bourdeau ym. 2019).

3.1.4 Opetus ja testaus

Datapohjaisessa mallinnusprosessissa datajoukko jaetaan tavallisesti kolmeen osaan: opetus-, testaus- ja validointijoukkoihin. Opetusdataa tarvitaan mallin opettamiseen. Validointivaiheen tarkoituksena on arvioida toteutettua algoritmia ja parantaa mallin sovitusta virittämällä mallin parametreja. Ylisovituksen välttämiseksi validointidatajoukon täytyy olla eri kuin opetusdata. Lopuksi on testausvaihe, jossa kehitetty algoritmi ajetaan aiemmin hyödyntämättömällä testidatalla, jolloin saadaan lopullinen arvio mallin parametreista ja sen suorituskyvystä. (Bishop 2006; Bourdeau ym. 2019)

Validointijoukko ja testijoukko menevät helposti sekaisin ja joissakin lähteissä validointijoukolla tarkoitetaan virheellisesti testijoukkoa. Ne on syytä kuitenkin erottaa. Aina validointijoukkoa ei käytetä, mutta se on yleensä suositeltavaa. Validointijoukkoa käytetään eri mallien suorituskvynarviointiin samalla, kun malliparametreja viritetään. Testijoukkoa käytetään lopullisen mallin arviointiin sen jälkeen, kun malli on kokonaan opetettu. Tarkoituksena on silloin testata lopullisen mallin suorituskvyyä ja yleistettävyyttä vasta mallin lopullisen määrittämisen jälkeen, toisin kuin validoinnissa.

3.1.5 Suorituskvyn arviointi

Mallin suorituskvyyä voidaan mitata suorituskvymittareilla. Niillä testataan ennustusalgoritmien luotettavuutta ja tarkkuutta. Niillä voidaan vertailla eri mallien paremmuutta ja tutkia virheen suuruutta mallinnetun ja mitatun arvon välillä. Usein käytettyjä tunnuslukuja mallin suorituskvyn vertailuun ovat keskimääräinen absoluuttinen virhe (MAE), keskimääräinen absoluuttinen suhteellinen virhe (MAPE), keskimääräinen neliövirhe (MSE), keskineliövirheen neliöjuuri (RMSE), RMSE:n vaihtelukerroin (CVRMSE), selitysaste (R^2), keskivirhe (MBE), normalisoitu keskivirhe (NMBE), keskimääräinen suhteellinen virhe (MRE) ja korrelaatio (r). (Amasyali & El-Gohary 2018):

$$MAE = \frac{1}{n} \sum_{i=1}^n |\hat{y}_i - y_i|, \quad (2)$$

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{\hat{y}_i - y_i}{y_i} \right| * 100 \%, \quad (3)$$

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2, \quad (4)$$

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2}, \quad (5)$$

$$\text{CVRMSE} = \sqrt{\frac{\frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2}{\bar{y}_i}} * 100 \%, \quad (6)$$

$$R^2 = 1 - \frac{\sum_{i=1}^n (\hat{y}_i - y_i)^2}{\sum_{i=1}^n (y_i - \bar{y}_i)^2}, \quad (7)$$

$$\text{MBE} = \frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i), \quad (8)$$

$$\text{NMBE} = \frac{\frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)}{\bar{y}_i} * 100 \%, \quad (9)$$

$$\text{MRE} = \left| \frac{\bar{\hat{y}} - \bar{y}_i}{\bar{y}_i} \right|, \quad (10)$$

$$r = \frac{\sum_{i=1}^n (y_i - \bar{y}_i)(y_i - \bar{\hat{y}}_i)}{ns_{y_i}s_{\bar{\hat{y}}_i}}, \quad (11)$$

missä \hat{y}_i on ennustettu ulostulon arvo ajanhetkellä i , y_i on todellinen ulostulon arvo ajanhetkellä i , \bar{y}_i on keskimääräinen mitattu ulostulon arvo, $\bar{\hat{y}}_i$ on ennustettujen arvojen keskiarvo, s_{y_i} ja $s_{\bar{\hat{y}}_i}$ ovat mitatun ulostulon ja ennusteen keskihajonnat, ja n on suorituskyvyn arvioinnissa käytettyjen datapisteiden määrä. (Bourdeau ym. 2019)

3.2 Datapohjaisten mallinnusmenetelmien rajoitteita

Datapohjaiset mallinnusmenetelmät eivät välttämättä suoriudu hyvin opetusdatan ulkopuolella. Mallin opetuksessa tehdyt oletukset eivät siten ole välttämättä yleistettävissä. Jos mallin opetukseen käytetään liian rajattua datajoukkoa, sen tarkkuus voi kärsiä. Mallinnukseen käytettävän datan täytyy siten olla tarpeeksi laaja ja sisältää informaatiota vaihtelevista olosuhteista esimerkiksi säätilassa, fysikaalisissa olosuhteissa tai operointitavoissa. Hyvin edustavan datajoukon kerääminen voi olla haastavaa, kallista ja työlästä. Toiseksi datapohjaiset mallit ovat black-box -malleja, joissa on omat rajoitteensa. Mallin tulomuuttujien ja lähdön välisiä suhteita voidaan kyllä ennustaa, mutta niiden vuorovaikutuksen taustalla olevia syitä voi olla vaikeaa ymmärtää. Tästä syystä parhaat menetelmät ovat usein niin sanottuja grey-box -malleja, jotka yhdistävät

black-box -mallin ja white-box mallin hyödyt ja minimoivat niiden haitat. (Amasyali & EL-Gohary 2018)

Mohanraj'n ym. (2012) mukaan neuroverkkolaskentaan liittyviä rajoituksia ovat ainakin yliopetus, ekstrapolointi ja verkon optimointiin liittyvät rajoitteet. Yliopetusta tapahtuu, kun mallin opetuskyky on liian korkea tai opetukseen käytetään liian monta iterointikertaa. Ylisovittunut malli kuvaa opetusdataa erittäin hyvin, mutta se ei ole luotettava todellisessa tilanteessa. Siksi on tärkeää optimoida opetuskertojen ja opetusdatan määrä. Ekstrapoloinnilla tarkoitetaan estimointia alkuperäisen opetusjoukon ulkopuolella. Verkon optimointi ja sen eri parametrien kokeilu yritysten ja erehdyksen kautta voi olla hyvin työläs ja aikaa vievä prosessi.

4 LÄMPÖKERTOIMEN ENNUSTAMISMENETELMIÄ

4.1 Keinotekoiset neuroverkot

Neuroverkko on melko yleistävä terminä. Se käsittää laajan joukon erityyppisiä ja eri tilanteisiin sopivia verkkorakenteita, joilla on eroavaisuuksia mallin tarkkuudessa ja monimutkaisuudessa. Tyypillisiä neuroverkkorakenteita ovat FFNN eli eteenpäin syöttävä neuroverkko (engl. Feed Forward Neural Network), BPNN eli takaisinkytketty neuroverkko (engl. Backpropagation Neural Network), radiaalinen kantafunktioverkko (engl. RBFNN, Radial Basis Function Neural Network), ANFIS eli adaptiivinen neurosumeaa päättelyjärjestelmä ja monikerroksinen perseptroni (engl. Multi-layer Perceptron, MLP). (Bourdeau ym. 2019)

Keinotekoisien neuroverkkojen etuna pidetään niiden nopeutta ja yksinkertaisuutta sekä sovellettavuutta monimuuttujaongelmien ja monimutkaisten riippuvuussuhteiden ymmärtämiseen muuttujien välillä. Niiden avulla on mahdollista selvittää epälineaarisia riippuvuussuhteita. Muihin menetelmiin verrattuna neuroverkkoja hyödyntämällä ei tarvita analyyttisiä yhtälöitä, sillä vaadittavan informaation oletetaan löytyvän opetusdatasta. Haluttu tulos voidaan saada myös rajallisesta opetusdatasta. Neuroverkot voivat toimia black-box -mallien tavoin, jolloin ei tarvita yksityiskohtaista tietoa mallinnettavasta kohteesta. (Mohanraj ym. 2012)

4.1.1 Verkon rakenne ja toiminta

Keinotekoiset neuroverkot pyrkivät jäljittelemään ihmisaivojen toimintaa esimerkeistä oppimalla. Verkon opetukseen käytetään opetusdataa. Verkon rakenne koostuu tulokerroksesta, lähtökerroksesta ja näiden välissä olevasta yhdestä tai useammasta piilokerroksesta. Kerrokset sisältävät neuroneita, jotka ottavat vastaan tietoa edellisiltä kerroksilta ja tuottavat lähdön aktivaatiofunktionsa perusteella. Tulokerros ottaa vastaan tulomuuttujadataa ja syöttää sitä verkossa eteenpäin. Lähtökerroksesta saadaan systeemin mallinnettu vaste. (Afram & Janabi-Sharifi 2015)

Piilokerros kerää ja käsittelee tuloista saatua tietoa ja lähettää muodostuneen vasteen verkossa eteenpäin. Jokaisessa piilokerroksessa lasketaan lähtö perustuen muuttujien ja aktivaatiofunktioden välisiin suhteisiin. Viimeinen piilokerros lähettää tiedon

lähtökerrokselle. Tällöin puhutaan eteenpäin syötetystä- eli feedforward-prosessoinnista. (Shin ym. 2019)

Tiedon prosessointi tapahtuu neuronien avulla. Tulokerroksessa on neuroneita yhtä paljon kuin tulomuuttujia. Lähtökerroksen neuronien määrä on sama kuin lähtöjen määrä. Piilokerroksessa on tyypillisesti neuroneita yhtä paljon tai enemmän kuin tulomuuttujia. Jokaisessa piilokerroksessa verkko oppii tulomuuttujista piirteitä ja muodostaa niistä oman vasteen. (Afram & Janabi-Sharifi 2015)

Neuroni koostuu synapseista, summaajasta ja aktivaatiofunktioista. Synapseilla tarkoitetaan neuronien välisiä kytkentöjä, joille määrätään oma painokerroin solmukohdan tärkeyden mukaan. Lähtö y ja seuraavan kerroksen tulosignaali voidaan määrittellä summauksen ja aktivaatiofunktion avulla seuraavasti:

$$y_i = f\left(\sum_{i=1}^m w_{il}x_i + b_l\right), \quad (12)$$

missä b_l on viive, w_{il} on painokerroin, x_i on tulosignaali, f on aktivaatiofunktio, joka on usein sigmoid-funktio:

$$f(x) = \frac{1}{1 + e^{-x}}. \quad (13)$$

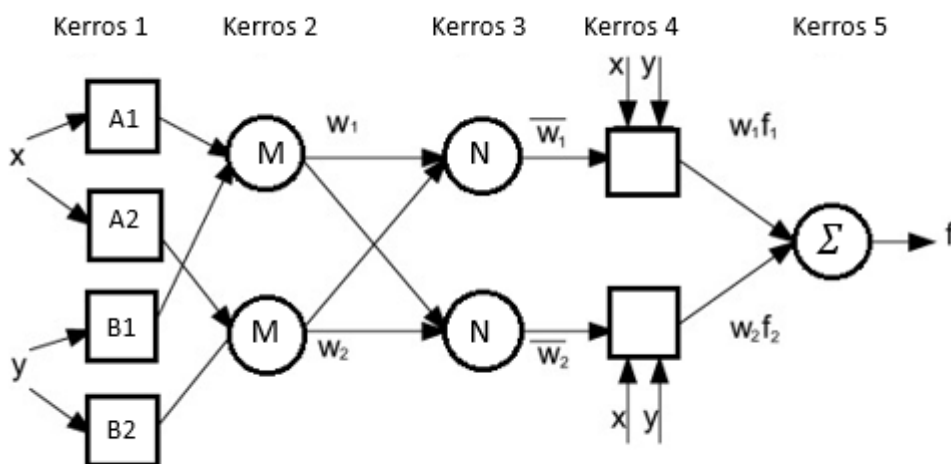
Aktivaatiofunktioiden kautta prosessoitu informaatio muutetaan vasteeksi. Sigmoid-funktion lisäksi aktivaatiofunktio voi olla myös kynnysfunktio (threshold) tai hyperbolinen tangenttifunktio (Shin ym. 2019).

Yksi neuroverkoissa laajasti käytetty rakenne on taaksepäin etenevä takaisinkytketty algoritmi ja sen eri muunnokset. Se optimoi alempien kerrosten painokertoimet lähdöstä saadun perusteella. Verkon vasteesta lasketaan virheet lähdön ja halutun lähdön välillä, jotka sen jälkeen ohjataan taaksepäin alempiin kerroksiin ja sen mukaan säädetään painokertoimia. Prosessia toistetaan, kunnes virhe on suositusarvon mukainen eli esimerkiksi 10^{-6} . (Mohanraj ym. 2012)

4.1.2 Adaptiivinen neurosumea päättelyjärjestelmä

Adaptiivinen neurosumea päättelyjärjestelmä eli ANFIS on monikerroksinen eteenpäin syöttävä verkko (engl. multilayer feedforward neural network, MLFFN), joka yhdistää neuroverkkojen oppimiskyvyn ja sumeaa logiikkaan perustuvan päättelyn ominaisuudet. Sumeilla systeemeillä voidaan pienentää optimoinnin hakuavaruutta tuomalla ennakkotietoja rajoitusten muodossa. Takaisinkytketyllä neuroverkkorakenteella taas voidaan automatisoida sumeiden parametrien viritykset.

ANFIS-verkon rakenne koostuu viidestä kerroksesta (Kuva 3). Ensimmäinen on sumea kerros, joka sisältää säädettäviä solmukohtia. Toisessa kerroksessa tuotetaan edellisen kerroksen tulos. Tämän kerroksen solmut ovat muuttumattomia. Toisesta kerroksesta lähtevät signaalit painotetaan arvoilla w_1 ja w_2 . Kolmannessa kerroksessa arvot normalisoidaan. Sen solmut ovat myös muuttumattomia. Neljännessä kerroksessa sumeat arvot selkeytetään. Kerroksen solmut ovat mukautuvia. Viimeisessä eli viidennessä kerroksessa lasketaan systeemin vaste summaamalla kaikki tulosisignaalit. (Mohanraj ym. 2012)



Kuva 3. ANFIS-verkon rakenne (mukaillen Mohanraj ym. 2012).

4.1.3 Radiaalinen kantafunktioverkko

Tässä neuroverkkorakenteessa käytetään aktivaatiofunktiona radiaalikantafunktiota. Verkko koostuu tavallisesti kolmesta kerroksesta: tulokerros, piilokerros ja lähtökerros.

Piilokerroksen yksiköt sisältävät RBF-funktion, joka on riippuvainen keskipisteen sijainnista ja leveydestä. Jokainen lähtö muodostuu piiloyksikköjen painotettuna summana. (Zhang ym. 2013)

4.1.4 Neuroverkkomalleihin perustuvat aiemmat tutkimukset

Esen ym. (2008a) ennustivat horisontaalisen maalämpöpumppusysteemin COP-arvoa neuroverkkomallilla käyttämällä tulomuuttujina lauhdutinyksikön tulo- ja lähtöilman lämpötiloja sekä jäätymisenestoaineen tulo- ja lähtölämpötiloja horisontaalisessa maalämmönvaihtimessa ja maaperän lämpötilaa. Esen ym. (2008b) ovat myös vertailleet neuroverkko- ja ANFIS-menetelmiä horisontaalisen maalämpöpumppusysteemin suorituskyvyn ennustamisessa käyttäen tulomuuttujina lauhduttimen tulo- ja lähtölämpötilaa ja maaperän lämpötilaa. ANFIS-menetelmä tuotti tässä tapauksessa tarkempia tuloksia.

Esen ym. (2008c) ennustivat minimidatajoukolla maalämpöpumppusysteemin päivittäistä lämpökerrointa ANFIS-menetelmällä, jossa käytettiin vastavirta-algoritmia ja sumeapainotettua esikäsittelyä (engl. Fuzzy Weighted Pre-processing, FWP). Muuttujina käytettiin lauhduttimen tulo- ja lähtöilman lämpötilaa, lämmönvaihtimen jäätymisenestoaineen tulo- ja lähtölämpötilaa sekä maaperän lämpötilaa.

Esen ja Inalli (2010) vertailivat ANN- ja ANFIS-menetelmiä vertikaalisen maalämpöpumppusysteemin lämpökertoimen ennustamisessa käyttäen tulomuuttujina höyrystimen ja lauhduttimen tuloilman ja lähtöilman lämpötilaa ja kylmäaineen tulo- ja lähtölämpötiloja. ANN-menetelmässä käytettiin MLP:tä ja opetusalgoritmina käytettiin kolmea eri algoritmia, joista nopeimmin ja tarkimmin suoriutui Levenberg-Marquardt-algoritmi. Lämmitystilassa kaikista paras malli saavutettiin ANFIS-rakenteella 'Triangular' kahdella jäsenyysastefunktiolla, jolloin RMSE- ja cov-arvot olivat 0,0238 ja 0,9276. Selitysaste oli puolestaan 0,9999, joka ainakin teoriassa on hyvin lupaava. Jäsenyysasteiden lisääminen johti mallin monimutkaistumiseen ja hitaampaan konvergoitumiseen.

Myös Sun ym. (2015) ovat vertailleet ANN- ja ANFIS-pohjaisia menetelmiä toimistorakennuksen pohjavesimaalämpöpumppusysteemin lämpökertoimen ennustamisessa. Tulomuuttujina käytettiin höyrystimen tulo- ja lähtöveden lämpötiloja,

lauhduttimen tulo- ja lähtöveden lämpötiloja ja kuorman puolen tulo- ja lähtölämpötiloja. Myös tässä tutkimuksessa molemmat menetelmät olivat luotettavia, mutta ANFIS-mallilla saatiin hieman parempia tuloksia ($RMSE = 0,096$) kuin neuroverkolla ($RMSE = 0,112$).

Zhang ym. (2013) käyttivät maalämpöpumpun lämpökertoimen ja vedenjakelun lämpötilan ennustamiseen radiaalista kantafunktioneuroverkkoa (engl. Radial Basis Function, RBFNN) käyttäen tulomuuttujina veden massavirtausta, rakennuksen lämpökuormaa ja maakierron veden tulolämpötilaa. Mallia käytettiin sen jälkeen energiaa säästävään ennustavaan säätöön käyttäen apuna APSO-algoritmia (engl. Adaptive Particle Swarm Optimization) tulomuuttujien optimointiin.

Fannou ym. (2014) käyttivät ANN-mallia DX-maalämpöpumppusysteemin (engl. Direct Expansion Geothermal Heat Pump) suorituskyvyn ennustamiseen. Tulomuuttujina käytettiin höyrystimen tulo- ja lähtölämpötilaa, höyrystimen tulo- ja lähtöpainetta, lauhduttimen jäähdytysveden tulolämpötilaa ja purkauspainetta. Lähtöinä saatiin lämpöpumpun suorituskykyyn vaikuttavat lämpökapasiteetti ja kompressorin sähkönkulutus. Paras malli saatiin käyttämällä Levenberg-Marquardt-optimointialgoritmia, kun piilokerros sisälsi 28 neuronaa.

Shin ym. (2019) kehittivät usean eri lämpöpumppuyksikön sisältävän maalämpöpumppusysteemin lämpökertoimen ANN-pohjaisen ennustusmallin neljällä tulomuuttujalla: Tulo- ja lähtölämpötila lämmönlähteestä sekä kuorman puolen tulo- ja lähtölämpötila. Ennustusmenetelmää sovellettiin sitten maalämpöpumppusysteemin reaaliaikaisen suorituskyvyn arviointiin, joka mahdollisti tehokkaamman operoinnin yhdessä kiertopumpun veden virtausnopeuden säädön kanssa. Tutkimuksessa ehdotettua operointitapaa käyttämällä voitiin tutkimuksen mukaan vähentää kiertoveden virtausnopeutta yli 29 % olemassa oleviin säätötapoihin verrattuna.

Neuroverkkoihin perustuvia tutkimuksia on tehty useita aiheeseen liittyen. Yhteenveto tutkimuksista on koottu Taulukkoon 1. Se sisältää menetelmien viittaukset, käytetyt menetelmät, tulo- ja lähtömuuttujat sekä mallien suorituskyvyn arviot.

Taulukko 1. Yhteenveto aiemmista neuroverkkopohjaisista tutkimuksista liittyen lämpökertoimen ennustamiseen.

| Lähde | Menetelmä | Tulomuuttujat | Systeemi ja lähtömuuttujat | Mallin suorituskyky (testiaineisto) |
|----------------------------|--------------|----------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Esen ym. 2008a | ANN | Lauhduttimen tulo- ja lähtölämpötilat, jäätymisenestoaineen tulo- ja lähtölämpötilat, maaperän lämpötila | Maalämpöpumppusysteemin COP (horisontaalinen lämmönvaihdin) | RMSE: 0,002 R ² : 0,9999 COV: 0,076 |
| Esen ym. 2008b | ANN ja ANFIS | Kompressorin tulo- ja lähtölämpötila, maaperän lämpötila | Maalämpöpumppu-systeemin COP | ANN: RMSE: 0,0100 R ² : 0,9999 COV: 0,2862 ANFIS: RMSE: 0,0047 R ² : 0,9999 COV: 0,1363 |
| Esen ym. 2008c | FWP-ANFIS | Lauhduttimen tulo- ja lähtöilman lämpötila, lämmönvaihtimien jäätymisenestoaineen tulo- ja lähtölämpötila, maaperän lämpötila | Maalämpöpumpun COP | RMSE: 0,0272 R ² : 0,9998 COV: 0,7733 |
| Esen ja Inalli 2010 | ANN ja ANFIS | Höyrystimen ja lauhduttimen tuloilman ja lähtöilman lämpötilat, vedenjäätymisenestoaineen tulo- ja lähtölämpötilat | Maalämpöpumpun COP (vertikaalinen lämmönvaihdin) | ANN: RMSE: 0,0344 R ² : 0,9998, COV: 1,3437 ANFIS: RMSE: 0,0238 R ² : 0,9999 COV: 0,9276 |
| Sun ym. 2015 | ANN ja ANFIS | Höyrystimen tulo- ja lähtöveden lämpötilat, lauhduttimen tulo- ja lähtöveden lämpötilat, kuorman puolen tulo- ja lähtölämpötilat | Maalämpöpumpun COP | ANFIS: RMSE: 0,096015, R ² : 0,9997 ANN: RMSE: 0,11227 R ² : 0,99959 |
| Zhang ym. 2013 | RBFNN | Rakennuksen kuorma, vedenkierron massavirtaus, maakierron tulolämpötila | Maalämpöpumppusysteemin COP ja käyttöveden lämpötila | COP: MRE: 1,75 %, RMSE: 0,132 R ² : 0,9996 Veden lämpötila: MRE: 4.53 %, RMSE: 0,150 R ² : 0,9985 |

| Lähde | Menetelmä | Tulomuuttujat | Systeemi ja lähtömuuttujat | Mallin suorituskyky (testiaineisto) |
|------------------------|-----------|---------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------|---------------------------------------------------------------------------------------|
| Fannou ym. 2014 | ANN | Höyrystimen tulo- ja lähtölämpötila, höyrystimen tulo- ja lähtöpaine, lauhduttimen jäähdytysveden tulolämpötila ja purkauspaine | Maalämpöpumpun lämpökapasiteetti ja kompressorin sähkönkulutus | RMSE: 0,16330 COV: 2,9319 R ² : 0,9991 |
| Shin ym. 2019 | ANN | Tulo- ja lähtölämpötila lämmönlähteestä, kuorman puolen tulo- ja lähtölämpötila | Useasta maalämpöpumpusta koostuvan systeemin COP | RMSE: 0,087579 R ² : 0,989475 *AV: [-0,14;0,11] *SV: [-0,06;0,03] |

*AV = absoluuttinen virhe, *SV = suhteellinen virhe

4.2 Satunnaismetsä

Satunnaismetsä eli Random Forest on Breimanin (2001) esittämä luokitteluun ja regressioon tarkoitettu ohjattu koneoppimismenetelmä, jossa mallin koulutuksessa käytetään useita eri päätöksentekopuita. Lähtönä saadaan parhaiten luokiteltu arvo tai keskimääräinen ennustus yksittäisiltä puilta. (Biau ja Scornet 2016)

RF-menetelmän tärkeimpiä hyötyjä on vähäinen ylisovittamisriski ja lyhyt opetusaika. Menetelmällä on myös mahdollista saada tarkkojakin ennustuksia ja se sopii menestyksekkäästi suurille datajoukoille.

Lu ym. (2019) käyttivät satunnaismetsä-menetelmää maalämpöpumppusysteemin suorituskyvyn ennustamiseen tuntitasolla. Menetelmää verrattiin takaisinkytketyn neuroverkkomallin kanssa ja satunnaismetsä osoittautui monessa suhteessa paremmaksi. Mallin robustisuus oli noin 3,3 % korkeampi kuin BPNN-mallilla, kun käytetty suorituskykymittari oli CVRMSE. Satunnaismetsä-mallin tulkittavuus oli myös parempi. Sen etuna oli myös sen yksinkertaisempi rakenne eikä se vaadi monimutkaista virittämistä. Lisäksi sillä oli huomattavasti lyhyemmät opetusajat.

4.3 Tukivektoregressio

Tukivektoregressio tukivektorikoneesta johdettu menetelmä, jossa sovitetaan hypertaso tai hypertasot, joilla datajoukko muunnetaan korkeampiin dimensioihin mallintamista varten. Menetelmän ominaisuus on, että se sallii virheen lisäämisen regressioon. Tämä virhe eli suurin marginaali on maksimietäisyys hypertason rajan ja lähimpien datapisteiden välillä. Näiden välistä etäisyyttä kutsutaan tukivektoreiksi.

Tässä menetelmässä käytetään Kernel-funktiota, joka voi olla lineaarinen, polynomi tai radiaalikantafunktio. Oikeanlaisen Kernel-funktion valinta on tärkeää mallin hyvän suorituskyvyn kannalta. SVR voi olla kohtuullisen hyvä ennustusmenetelmä muihin verrattuna. Sen haasteena on se, että parametrien optimointi on vaikeaa, mutta oleellista ennustustarkkuuden kannalta. Kernel-funktion tarkempi määrittäminen voi olla myös haastavaa. (Bourdeau ym. 2019)

Wang ym. (2019) ennustivat esilämmittimen sisältävän höyryinjektioon perustuvan lämpöpumpun lämpökapasiteettia ja lämpökerrointa käyttämällä tukivektoregressiota (SVR). Tutkimuksessa paras tulomuuttujajoukko saatiin seuraavilla piirteillä ja muuttujilla: ulkolämpötila, sisälämpötila, sisäilman massavirtaustaso, kompressorin nopeus, injektiopaine, latauspaine sekä näiden paineiden suhde. Ennustusmallin maksimivirhe lämpökapasiteetille oli 8,25 % ja lämpökertoimelle 8,33 %.

Esen ym. (2008d) käyttivät regressioon perustuvaa Lin-Kernel SVM-menetelmää maalämpöpumpun lämpökertoimen ennustamisessa saavuttaen sillä paremman tuloksen kuin ANN- ja ANFIS-menetelmällä. SVM oli myös huomattavasti nopeampi johtuen yksinkertaisemmasta rakenteesta.

Tukivektorikoneeseen pohjautuvia menetelmiä ovat hyödyntäneet myös Markovic ym. (2017) ilma-vesilämpöpumpun kuuman veden kulutusprofiilin ja lämpökertoimen ennustamiseen viikkotasolla. Mallin tulomuuttujina käytettiin asukkaan käyttäytymiseen liittyviä tekijöitä eli sisäilman lämpötiloja sekä kotitalouden ja lämpöpumpun energiankulutusta. MAE oli 0,238 ja sen keskihajonta 0,492.

4.4 Muita menetelmiä

Yksinkertaisimmillaan lämpökerroin saadaan estimoitua yksinkertaisilla lineaarisilla tai epälineaarisilla matemaattisilla kaavoilla ja polynomifunktioilla, joita voidaan simuloida esimerkiksi laitevalmistajan tuottaman datan avulla. Tällaisten mallien käyttö on melko yleistä energiaoptimointiongelman yhteydessä ja rakennusten lämmityksen säädön osana.

Lämpöpumppujen yleinen säätöstrategia on olettaa lämpökertoimen vakioksi optimointiongelman yksinkertaistamisen takia. Tarkkuus on kuitenkin silloin yleensä vaatimaton eikä mahdollisia kuormitusvaihteluita voida arvioida. Epälineaarisella yhtälönsovituksella tarkkuutta voidaan lisätä, mutta mallista tulee siten monimutkaisempi (Cupeiro Figueroa ym. 2020). Esimerkiksi Verhelstin ym. (2012) tutkimuksessa rakennuksen energiankulutuksen optimointiin sovelletussa malliprediktiivisessä säädössä lämpöpumpun lämpökerroin mallinnettiin vakiona riippumatta ulkolämpötilasta, käyttöveden lämpötilasta ja muista vaikuttavista tekijöistä.

Zou ja Xie (2017) käyttivät yksinkertaista matemaattista lämpökerroinmallia liittyen lämpöpumppuun, joka ottaa lämmön järvivedestä. Lämpökerroinmallissa käytettiin järviveden lämpötilaa, joka ennustettiin säädään perustuvalla mallilla sekä keskimääräistä käyttäjän puolen lähtevän veden lämpötilaa. Mitattu ja ennustettu COP korreloivat keskenään niiden eron ollessa viilennyskaudella 2,103 % ja lämmityskaudella 3,670 %. Työssä tutkittiin kausittaista lämpökertoimen vaihtelua.

Rastegarpour ym. (2019) tutkivat erilaisia säätöön soveltuvia ilma-vesilämpöpumpun COP-ennustusmalleja, jotka olivat käytännössä neliöllisiä tai lineaarisia polynomifunktioita, joissa huomioitiin lämpökertoimeen vaikuttavia tekijöitä. Paras ennustus saatiin monimutkaisimmalla neliöllisellä kaavalla, joka koostui kolmesta tulosta, joita olivat ulkolämpötila, normalisoitu kompressorin teho ja lauhduttimeen tulevan veden lämpötila. Lämpökerroin on mallinnettu siinä staattisena funktiona, joka esittää tietyn aikavälin keskimääräisen COP-arvon. Ferrarini ym. (2020) arvioivat, että edellisen kaltaisen mallin soveltaminen käytäntöön ei ole yksinkertaista, koska se vaatisi useita sensoreita ja monimutkaista systeemin käyttäytymisen mallinnusta sekä toimivaa säädintä. Lämpöpumpun suorituskyvyn optimoinnin tavoitteena on, että saadaan tarkasti mallinnettua sen käyttäytymistä lyhyellä aikavälillä eli tuntitasolla. Optimoinnissa tulisi ottaa huomioon myös sääolosuhteet ja sähkökuormitus.

Mallin yksinkertaisuudesta johtuvan ennustusvirheen kompensoimiseksi Rastegarpour ym. (2021) kehittivät ilma-vesilämpöpumpulle COP-ennustusmallin mukautuvalla laajennetulla Kalman-suotimella, jota sovellettiin malliprediktiiiviseen säätöön. Mallin rakentamisessa käytettiin ulkolämpötilaa ja kuormaprofiilia. Malli kykeni kuvaamaan suhteellisen hyvin epälineaarisia ajassa muuttuvia lämpöpumpun tehokkuuteen vaikuttavia ominaisuuksia. Kehitetyllä säätöratkaisulla kyettiin parantamaan lämpöpumpun energiatehokkuutta. Verrattuna klassisempiin säätöratkaisuihin, joissa COP on mallinnettu keskimääräisinä arvoina tai vakiona, tällaisella mukautuvalla ratkaisulla saavutettiin parempi suorituskky, joka johti merkittäviin energiasäästöihin.

Lineaarista monimuuttujaregressiomallia (engl. Multiple Linear Regression, MLR) maalämpöpumpun suorituskvyn ennustamisessa ovat käyttäneet esimerkiksi Simon ym. (2016). Kyseessä oli yksinkertainen toisen kertaluokan lineaarinen MLR-malli, joka perustui laitevalmistajan dataan liittyen nesteiden lämpötiloihin ja virtausnopeuksiin. Lähtöinä saatiin lämpökapasiteetti ja COP. Mallin sovitukseen käytettiin kahdeksaa tilastollisesti merkittävää muuttujaa, jotka identifioitiin 36:n eri havainnon joukosta. Lämpökertoimen ennustusmallin ennustusvirhe oli 2,6 %, 4,9 % ja 3,2 % kolmelle eri maalämpöpumpulle. Lähteen mukaan maalämpöpumppusysteemin globaalia käyttäytymistä voidaan ennustaa MLR-mallilla rajallisellakin datalla hyväksyttävissä olevalla tarkkuudella.

Taulukkoon 2 on koottu yhteenveto luvuissa 4.2–4.4 mainituista lämpökertoimen ennustamiseen liittyvistä tutkimuksista.

Taulukko 2. Yhteenveto RF, SVR ja muista mallinnusmenetelmistä lämpökertoimen ennustamiseen.

| Lähde | Menetelmä | Tulomuuttujat | Systeemi ja lähtömuuttujat | Mallin suorituskyky |
|------------------------------|---------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------|
| Lu ym. 2019 | RF | Maaperän lämpötila, vesipumpun imu- ja purkupaineet, tulevat ja lähtevät veden lämpötilat maalämmönvaihtimessa ja käyttäjäpuolella, tilavuusvirtausnopeus, kuormasuhde | Maalämpöpumppu-systeemin pitkän aikavälin COP | R^2 : 0,89–0,97 MAPE: 1,15–1,93 RMSE: 0,07–0,12 |
| Wang ym. 2019 | SVR | Kompressorin pyörimisnopeus, ulkolämpötila, sisälämpötila, massavirtausnopeus, purkutodennäköisyys, injektion todennäköisyys, injektion painesuhde | Höyryinjektiolämpöpumpun lämpökapasiteetti ja COP | Maksimivirhe: 8,33 % (COP) |
| Esen ym. 2008cd | Lin-kernel SVM | Lauhduttimeen saapuvan ja lähtevän ilman lämpötila, kylmäaineen lämpötila maalämmönvaihtimessa, maaperän lämpötila | Maalämpöpumppu-systeemin COP | RMSE: 0,0027 R^2 : 0,9999 COV: 0,0773 |
| Markovic ym. 2017 | SVM | Ilman lämpötilat eri huoneissa ja ulkona, kotitalouden ja lämpöpumpun energiankulutus | Ilma-vesilämpöpumpun lämpimän veden kulutus ja COP | MAE: 0,238 |
| Zou & Xie 2017 | Matemaattisesti simuloitu | Lauhduttimeen tulevan veden lämpötila, höyrytimestä lähtevän veden lämpötila, mallin kertoimet | Vedestä lämmön ottavan lämpöpumpun COP | R^2 : 0,964 |
| Rastegarpour ym. 2019 | Neliölliset ja lineaariset polynomifunktiot | Ulkolämpötila, veden lämpötila, kompressorin tehontaaajuusprosentti | Ilma-vesilämpöpumpun COP | NRMSE: 0,5 % |
| Rastegarpour ym. 2021 | Lineaarinen ensimmäisen kertaluokan polynomimalli Kalman-suotimella | Ulkolämpötila ja paluuveden lämpötila vesisäiliöstä | Ilma-vesilämpöpumpun COP_1 (ottaa huomioon kuormaprofiilin) ja COP_2 (vain ulkolämpötila huomioon) | Ennustusvirhe COP_1 : 4 % suotimella, 16 % ilman COP_2 : 13 % suotimella, 35 % ilman |
| Simon ym. 2016 | MLR | Fluidien lämpötilat ja virtausnopeudet | Kolmen eri maalämpöpumpun COP | R^2 : 0,998, 0,999, 0,999 Ennustusvirheet (CV): 2,62 %, 4,91 % ja 3,21 % RMSE: 0,129; 0,237; 0,156 |

4.5 Yhteenveto menetelmistä lämpökertoimen ennustamisessa

Lämpöpumppujen lämpökertoimen ennustamiseen liittyviä tutkimuksia löytyy selvästi eniten maalämpöpumppuihin liittyen. Maalämpöpumppuilla on tärkeä rooli rakennusten energiatehokkuuden parantamisessa ja lämpökerroin on oleellinen tekijä niiden suorituskvyyllle. Maalämpöpumppu on myös investointina kallis, joten on ymmärrettävää, että sen suorituskvyyä halutaan parantaa. Melko paljon löytyi tutkimustietoa myös ilmalämpöpumppujen suorituskvyyistä (Markovic ym. 2017; Rastegarpour ym. 2019; Rastegarpour ym. 2021).

Parhaan menetelmän valinta ei ole aina kovin yksinkertaista. Jokaisella menetelmällä on omat hyvät ja huonot puolensa. Mallin onnistumiseen vaikuttavat monet asiat, kuten datan laatu, sovelluskohde ja mallin parametrien optimointi. Älykkäillä menetelmillä on mahdollista saada hyvin tarkkojakin ennustuksia, mutta toisaalta ne voivat olla monimutkaisia ja laskenta-ajat pitkiä. Joskus yksinkertainen malli on parhaiten sovellettavissa ja yleistettävissä, vaikka jäisi tarkkuudessa muista menetelmistä jälkeen. Monet säädössä käytetyt COP-ennustusmallit perustuvatkin melko yksinkertaisiin lineaarisiin tai epälineaarisiin mallirakenteisiin.

Suorituskvyyt mittarit antavat jonkinlaista suuntaa mallin suorituskvyyistä. Ne eivät kerro koko totuutta tai yksiselitteisesti sitä, että mikä malli on paras. Suorituskvyyt mittarit voivat olla erinomaiset, mutta siitä huolimatta esimerkiksi mallin yleistettävyyks voi olla huono. Siihen vaikuttaa myös muut mallinnuksen osatekijät. Mallien suorituskvyyä on mitattu esimerkkitapauksissa useimmiten neliöllisellä keskiarvolla RMS ja sen virheellä RMSE, selityssasteella R^2 , variaatiokertoimella COV ja suhteellisella keskivirheellä MRE. Mallinnus on onnistunut, jos ennustetut arvot vastaavat hyvin mitattuja arvoja ja niiden välinen ero on pieni. Mitä lähempänä nollaa RMS, RMSE, COV ja MRE ovat, sitä parempi malli. Selityssaste R^2 puolestaan voi vaihdella nollan ja ykkösen välillä. Mitä lähempänä se on ykköstä, sitä lähempänä ennustetut arvot ovat mitattuja arvoja. Selityssaste lähentelee useimmissa neuroverkkopohjaisissa esimerkkitapauksissa ykköstä. Ennustetut arvot ovat siis hyvin lähellä mitattuja arvoja, mikä sinällään on toivottavaa. Tästä herää kuitenkin kysymys, onko käytetty malli todellisuudessa niin hyvä, vai johtuuko lupaava suorituskvyy esimerkiksi ylikouluttamisesta. Ylikoulutus on neuroverkkomenetelmillä yleinen haaste. Neuroverkkomenetelmät ovat myös hyvin riippuvaisia datan laadusta, mikä luonnollisesti vaikuttaa mallin luotettavuuteen.

Aiempien tutkimusten perusteella voidaan todeta, että useimmat datapohjaiset lämpöpumppujen lämpökerroinmallit perustuvat älykkäisiin menetelmiin, kuten neuroverkkopohjaisiin malleihin ja tukivektoregressioon (SVR). Myös satunnaismetsä-menetelmää on käytetty lämpöpumpun lämpökertoimen ennustamisessa lupaavin tuloksin (Lu ym. 2019). Siihen liittyviä tutkimuksia lämpökertoimen ennustamisesta ei löydy kuitenkaan kovin montaa. Maalämpöpumppusysteemi on monimuuttujainen systeemi, jota on vaikea mallintaa perinteisillä menetelmillä (Esen ym. 2008d). Oletettavasti tästä syystä tutkimus painottuu älykkäisiin menetelmiin, kuten taulukkojen 1 ja 2 yhteenvedosta nähdään.

ANN on hyvin suosittu menetelmä mallintamaan monimutkaisia systeemejä, joita voi olla vaikeaa mallintaa perinteisillä menetelmillä. Niiden etuna on hyvä oppimiskyky ja sopivuus epälineaarisiin systeemeihin, jonka takia neuroverkkomalleilla on runsaasti sovelluksia rakennusten energiapohjajärjestelmien ennustuksessa. Yksi lupaava menetelmä niin hyvien tarkkuusmittareiden kuin käytäntöön soveltamisen osalta on Shinin ym. (2019) kehittämä ANN-pohjainen maalämpöpumppusysteemin COP-ennustusmalli, jolla pystyttiin ennustamaan lämpökerroin reaaliajassa operoinnin aikana. Ennustusmallin ja kehitetyn virtausnopeussäätimen yhdistelmällä saatiin lähteen mukaan noin 23 % energiasäästöt.

Fannou ym. (2014) ja Esen ja Inalli (2010) osoittivat omissa tutkimuksissaan, että BPNN-pohjaisilla neuroverkoilla paras toteutus saatiin käyttämällä Levenberg-Marquardt -algoritmia sen parametrien optimointiin. Neuroverkko- ja ANFIS malleja on vertailtu useammassa tutkimuksessa (Esen ym. 2008b; Esen ja Inalli 2010; Sun ym. 2015) ja neurosumea ANFIS-menetelmä on tuottanut niissä kaikissa hieman parempia tuloksia.

Kuitenkin Esenin ym. (2008d) mukaan neuroverkko ja ANFIS-malleja vielä paremmin suoriutuu tukivektorkoneeseen pohjautuva lineaarinen regressiomenetelmä (Lin-kernel SVR). Lähteen mukaan SVR-malli on paremmin yleistettävä ja robustimpi kuin neuroverkko- ja ANFIS-mallit. Menetelmä myös konvergoituu globaaliin optimiin helpommin. SVR tuotti pienemmän RMS-arvon ja COV-arvon, joka viittaa parempaan suorituskyykyyn. SVR-menetelmää on sovellettu niin maalämpöpumppujen (Esen ym. 2008d), ilmavesilämpöpumppujen (Markovic ym. 2017) kuin höyryinjektioon perustuvan lämpöpumpun lämpökertoimen ennustamiseen (Wang ym. 2019), joten sovelluskohteita on useita.

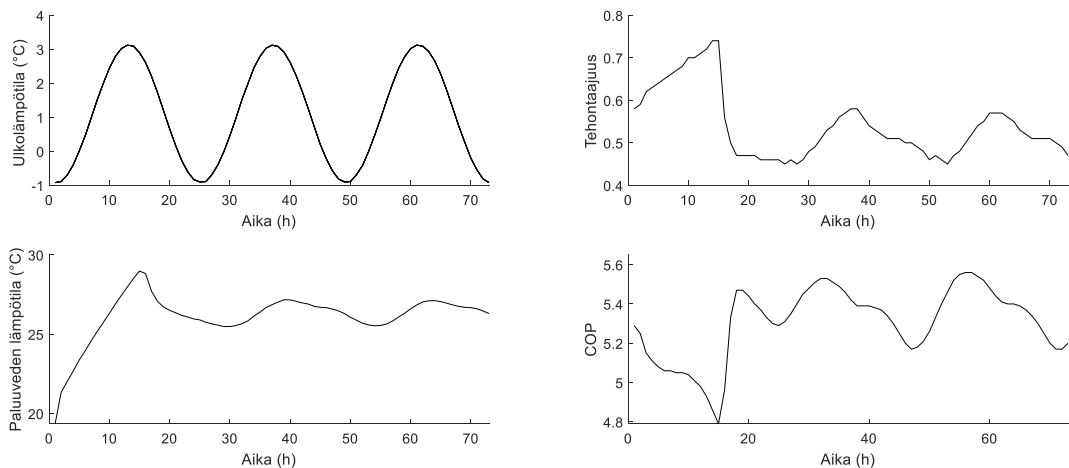
Muista tutkituista menetelmistä potentiaalisimmalta menetelmältä vaikuttaa Rastegarpourin ym. (2021) käyttämä COP-ennustusmalli, jossa käytettiin mukautuvaa laajennettua Kalman-suodinta, jonka avulla yksinkertaistakin lineaarista mallia käyttämällä saatiin luotettavia ennusteita. Menetelmää käytettiin myös onnistuneesti ilmavesilämpöpumpun MPC-säädössä ja se tuotti selviä energiasäästöjä, joten vastaaville ratkaisuille voi löytyä käyttöpotentiaalia Suomessakin.

5 LÄMPÖKERTOIMEN MALLINNUS JA SIMULOINTI

Kokeellinen osa käsittelee dynaamisten regressiomallien ja neuroverkkomallien käytännön soveltamista lämpökertoimen ennustamisessa ja mallinnusprosessin eri vaiheista, kuten mittausaineiston käsittelystä, mallin rakentamisesta ja tuloksista. Kaikki datankäsittely-, mallinnus- ja simulointityö tehtiin MATLAB®-ohjelmistolla.

5.1 Mittausaineistot

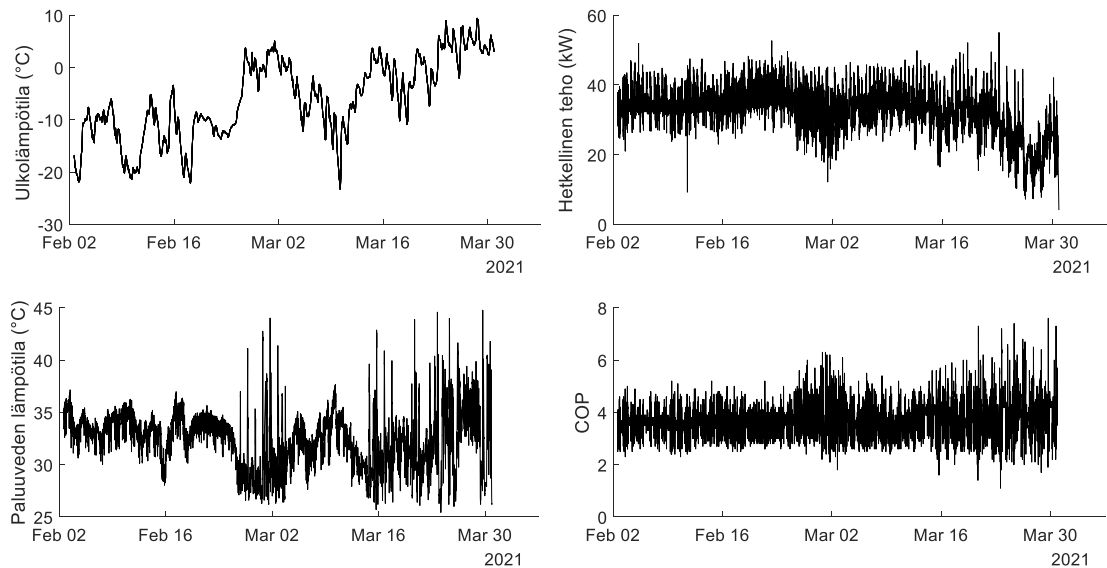
Käytössä oli kaksi datajoukkoa. Mittausaineisto 1 oli simuloitua tuntitason dataa, joka on johdettu (Rastegarpour ym. 2019) käyttämistä regressiomalleista. Se sisälsi 73 datapistettä jokaista muuttujaa kohden tunnin välein. Mittausaineisto 2 oli eräästä poistoilmalämpöpumpusta peräisin olevaa mitattua dataa, joka kerättiin HögforsGST Oy:n toimittaman Fiksu Control System -järjestelmästä, joka toimii erään Oulussa sijaitsevan taloyhtiön lämmityksen ohjausjärjestelmänä. Molemmat datajoukot sisälsivät samat tulomuuttujat (ulkolämpötila, paluuveden lämpötila ja kompressorin teho), mutta mittausaineisto 1:ssä hetkellisen tehon tilalta oli kompressorin tehontaajuusprosentti. Mittausaineisto 1 on esitetty kuvassa 4.



Kuva 4. Mittausaineiston 1 simuloitua ulkolämpötila, kompressorin tehontaajuus, paluuveden lämpötila ja COP.

Mittausaineistossa 2 oli dataa kahden kuukauden ajalta vuoden 2021 helmikuun alusta maaliskuun loppuun. Se sisälsi ulkolämpötilan, latauspiirin paluulämpötilan,

kompressorin hetkellisen tehon sekä COP-arvon. Alkuperäisessä datassa ulkolämpötilamittaus oli 10 minuutin välein ja muut mittaukset yhden minuutin resoluutiolla. Työssä käytetty esikäsitelty mittausaineisto 2 on esitetty kuvassa 5.



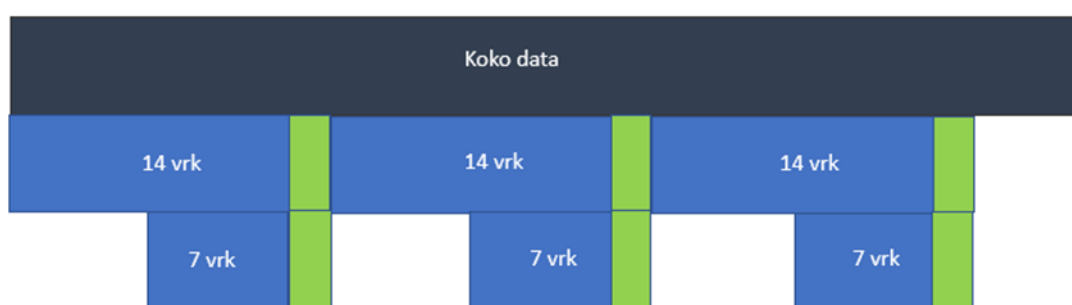
Kuva 5. Mittausaineiston 2 mitattu ulkolämpötila, kompressorin hetkellinen teho, latauspiirin paluuvien lämpötila ja lämpökerroin.

Mittausaineiston 2 ulkolämpötila sisälsi muutamia puuttuvia arvoja, jotka korvattiin muuttujan edellisellä arvolla. Useampia datapisteitä sisältävät kompressorin hetkellinen teho, paluulämpötila ja COP vaativat enemmän käsittelyä, sillä puuttuvien arvojen lisäksi mittausaineisto sisälsi myös joitakin selkeästi muista poikkeavia arvoja. Merkittävästi poikkeavia arvoja arvioitiin lineaarisen interpoloinnin avulla. Poikkeamien määrittämiseen käytettiin liikkuvaa mediaania, kun kynnyskriteerinä oli teholle 7, paluuviedelle 6 ja referenssi COP-arvolle 6. Kynnysarvot määrittivät, kuinka paljon arvon piti poiketa mediaanista, että se tulkittiin poikkeamaksi.

Oleellinen vaihe datan esikäsittelyssä oli datajoukkojen muuttaminen täsmälleen samaan aikaikkunaan ja samanpituisiksi vektoreiksi siten, että ajasta riippuvat muuttujat ovat keskenään vertailukelpoisia. Mittausaineistolle 2 datan aikaikkunaksi muokattiin 1.2.2021 klo 13:46 – 31.3.2021 klo 21:00. Data saatiin tasaväliseksi MATLAB®:n *retime*-funktioilla, joka perustui myös lineaariseen interpolointiin. Tällöin mallinnusta varten kullekin muuttujalle saatiin datapisteitä yhteensä 81075, yhden minuutin näyteväliä.

Mittausaineisto 1 oli suppea, joten se jaettiin yhteen 45 tunnin opetusjoukkoon ja 26 tunnin testiosaan. Mittausaineistoa 1 käytettiin vain regressiomallien kanssa.

Mittausaineisto 2 jaettiin kolmeen yhtä pitkään osaan. Jokaisesta osasta käytettiin ensimmäiset 14 vrk mallien parametrien identifiointiin ja sitä seuraavat 2 vrk mallien testaamiseen. Kun opetusjakson pituus oli 7 vrk, käytettiin täsmälleen samoja testijoukkoja, mutta testijaksoa ennen oleva opetusjoukko puolittui. Kuvassa 6 on esitetty datan jakoperiaate. Aikariippuvuudesta johtuen testijakso täytyi valita aina heti opetusjoukon jälkeen.



Kuva 6. Periaatekuva mittausaineiston 2 jaosta. Sinisellä opetusjaksot ja vihreällä kahden vuorokauden testijaksot.

5.2 Lämpökertoimen ennustamiseen sovelletut mallirakenteet

5.2.1 Regressiomallit

Työssä testattiin yhteensä viittä erilaista regressiomallia. Näistä kaksi olivat lineaarisia ja neljä neliöllisiä. Yksinkertaisimmassa mallissa (COPL1) lämpökertoimen ennustaminen perustui ainoastaan ulkolämpötilan lineaariseen sovitukseen ja monimutkaisin malli (COPQ3) sisälsi ulkolämpötilan lisäksi latauspiirin paluulämpötilan sekä kompressorin hetkellisen tehon ja niiden neliölliset arvot. Regressiomalleja testattiin molemmilla datajoukoilla. Ennustushorisontti oli kaksi vuorokautta.

Yhtälöiden 14 ja 15 mukaiset mallit on esitetty lähteessä Rastegarpour ym. (2021). Työssä toteutettujen regressiomallien perustana ovat Rastegarpour ym. (2019) käyttämät polynomimallit, jotka muunnettiin dynaamiseen muotoon. Heillä oli käytössä ilma-vesilämpöpumpun dataa, kun taas tässä työssä data oli peräisin

poistoilmalämpöpumpusta. Toinen ero oli, että he käyttivät hetkellisen tehon tilalta kompressorin tehontaajuusprosenttia. Regressiomallien dynaamiset yhtälöt ovat siten:

$$COP_{L1}(t) = COP_{L1}(t-1) + a_1(-COP_{L1}(t-1) + b_1 + c_1 T_{out}(t)), \quad (14)$$

$$COP_{L2}(t) = COP_{L2}(t-1) + a_2(-COP_{L2}(t-1) + b_2 + c_2 T_{out}(t) + d_2 T_B(t)), \quad (15)$$

$$COP_{Q2}(t) = COP_{Q2}(t-1) + a_3(-COP_{Q2}(t-1) + b_3 + c_3 T_{out}(t) + d_3 T_B(t) + e_3 T_{out}(t)^2 + f_3 T_B(t)^2 + g_3 T_{out}(t) T_B(t)), \quad (16)$$

$$COP_{Q3}(t) = COP_{Q3}(t-1) + a_4(-COP_{Q3}(t-1) + b_4 + c_4 T_{out}(t) + d_4 T_B(t) + e_4 P(t) + f_4 T_{out}(t)^2 + g_4 T_B(t)^2 + h_4 P(t)^2 + i_4 T_{out}(t) T_B(t) + j_4 T_B(t) P(t) + k_4 T_{out}(t) P(t)), \quad (17)$$

$$COP_{Q4}(t) = COP_{Q4}(t-1) + a_5(-COP_{Q4}(t-1) + b_5 + c_5 T_{out}(t) + d_5 P(t) + e_5 T_{out}(t)^2 + f_5 P(t)^2 + g_5 T_{out}(t) P(t)), \quad (18)$$

missä $a-k$ ovat mallin parametreja, T_{out} on ulkolämpötila, P on kompressorin hetkellinen teho ja T_B on latauspiirin paluuveden lämpötila. Mittausaineistolla 2 ajettujen regressiomallien parametrien $a-k$ keskimääräiset arvot keskihajontoineen on esitetty taulukossa 3.

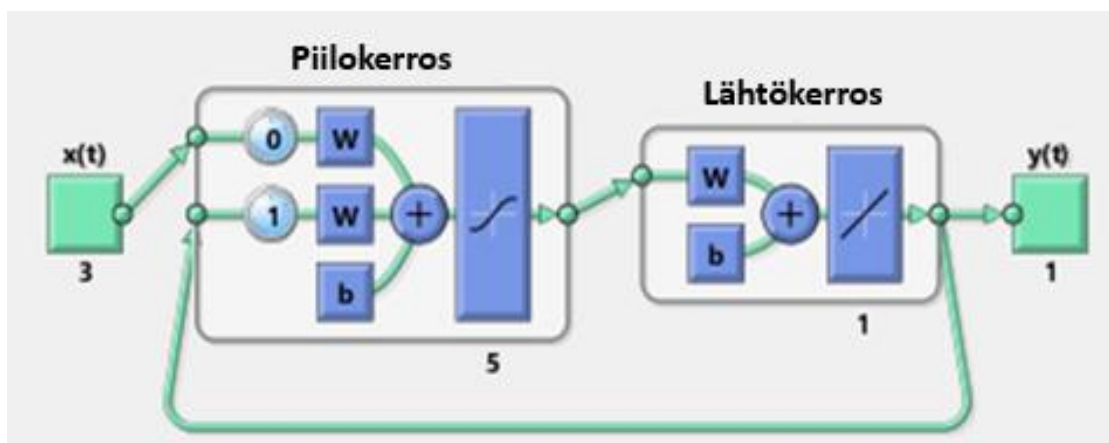
Taulukko 3. Dynaamisten regressiomallien parametrien $a-k$ keskimääräiset arvot sekä niiden keskihajonnat mittausaineistolla 2.

| Parametri | Malli | | | | |
|-----------|------------|-------------|-------------|-------------|-------------|
| | L1 | L2 | Q2 | Q3 | Q4 |
| a_{1-5} | 0,177±0,16 | 1,142±0,45 | 1,019±0,36 | 1,134±0,31 | 1,419±0,20 |
| b_{1-5} | 3,896±0,15 | 10,504±2,29 | 17,860±4,14 | 13,270±6,48 | 8,714±1,05 |
| c_{1-5} | 0,027±0,10 | -0,029±0,02 | -0,238±0,20 | -0,197±0,37 | 0,038±0,02 |
| d_{2-5} | | -0,218±0,07 | -0,734±0,39 | -0,065±0,64 | -0,227±0,04 |
| e_{3-5} | | | 0,002±0,00 | -0,329±0,09 | 0,000±0,00 |
| f_{3-5} | | | 0,009±0,01 | 0,001±0,00 | 0,002±0,00 |
| g_{3-5} | | | 0,009±0,01 | -0,008±0,02 | -0,001±0,00 |
| h_4 | | | | -0,001±0,00 | |
| i_4 | | | | 0,004±0,02 | |
| j_4 | | | | 0,012±0,01 | |
| k_4 | | | | 0,001±0,00 | |

5.2.2 Epälineaariset autoregressiiviset neuroverkkomallit

Työssä toteutettiin kolme epälineaarista autoregressiivistä neuroverkkomallia (NARX) käyttämällä mitattua dataa (Mittausaineisto 2). Mallirakennetta voidaan käyttää lähtömuuttujan (lämpökerroin) aikasarjan ennustamiseen sen aiempien arvojen avulla. Mallien toimintaperiaate on sama, mutta niissä on testattu eri tulomuuttujakombinaatioita. NARX1 sisälsi yhden tulomuuttujan (ulkolämpötila), NARX2 kaksi tulomuuttujaa (ulkolämpötila ja kompressorin hetkellinen teho) ja NARX3-mallirakenne kolme tulomuuttujaa (edellä mainittujen lisäksi latauspiirin paluulämpötila).

Ensimmäisenä määritettiin neuroverkon rakenne MATLAB[®]:n *narxnet* -funktiolla, jonka jälkeen verkko koulutettiin opetusdatan avulla. Tulomuuttujien viive oli nolla ja ulostulon viive oli yksi. Piilokerrosten lukumäärä oli yksi. Opetusalgoritmina käytettiin Levenberg-Marquardt -menetelmää. Opetuksen jälkeen malli validoitiin testidatalla. Kuvassa 7 on esitetty NARX3-neuroverkkomallin rakenne, jossa vasemmalla on tulomuuttujat, sen jälkeen tulee piilokerros, joka sisältää tulon ja ulostulon viiveet sekä neuronien lukumäärän. Lähtönä saadaan lämpökertoimen kahden vuorokauden ennuste. Verkko on suljettu, jolloin se käyttää edellisiä lähdön arvoja seuraavaan ennusteeseen. Malleja testattiin eri neuronimäärillä.

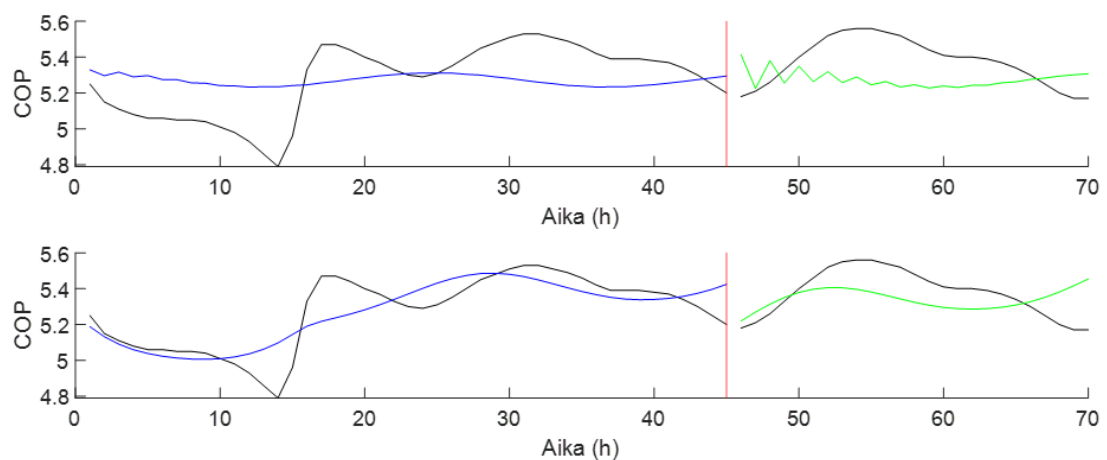


Kuva 7. Neuroverkkomallin NARX3 rakenne, joka sisältää yhden piilokerroksen ja lähtökerroksen. (mukaillen MATLAB[®])

5.3 Tulokset ja niiden tarkastelu

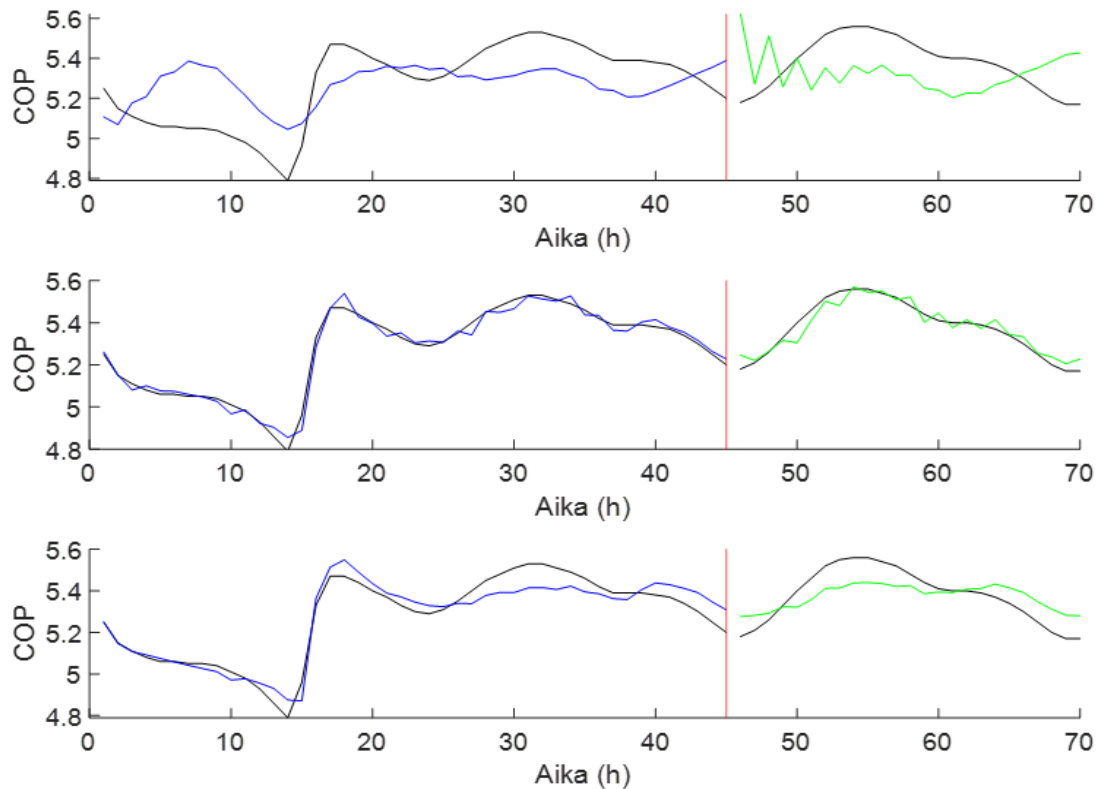
5.3.1 Regressiomallit: simuloitu data

Mittausaineistolla 1 saatiin kuvassa 8 olevat tulokset lineaaristen mallien L1 ja L2 opetus- ja testidatalla. Kuvasta nähdään, että L1-mallin 26:n tunnin ennuste on epätarkka, sillä se ei näytä seuraavan referenssiarvoa. L2-malli on sen sijaan huomattavasti tarkempi. Mallin L1 testijakson RMSE oli 0,181 ja MAPE oli 3,023 % ja L2:n vastaavasti oli 0,144 ja 2,27 %.



Kuva 8. Opetusjakson (sininen) ja ennusteen validointitulokset (vihreä) dynaamisille lineaarisille regressiomalleille L1 (ylempi) ja L2 (alempi) mittausaineiston 1 testidatalla. Musta viiva on lämpökertoimen simuloitu referenssiarvo.

Kuvassa 9 on esitetty simulointitulokset regressiomalleille Q2, Q3 ja Q4 mittausaineistolla 1. Kuvasta 9 nähdään, että Q3 on selkeästi tarkin malli ja Q4 myös seuraa hyvin referenssiarvoa. Mallin Q3 testiosan RMSE oli 0,038 ja MAPE 0,570 % ja mallin Q4 vastaavat olivat 0,086 ja 1,422 %. Sen sijaan Q2 ennusti yllättävän huonosti tällä datalla, sillä ennuste ei juuri korreloinut referenssiarvon kanssa. Sen RMSE testiosalla oli 0.197 ja MAPE 3,02 %.

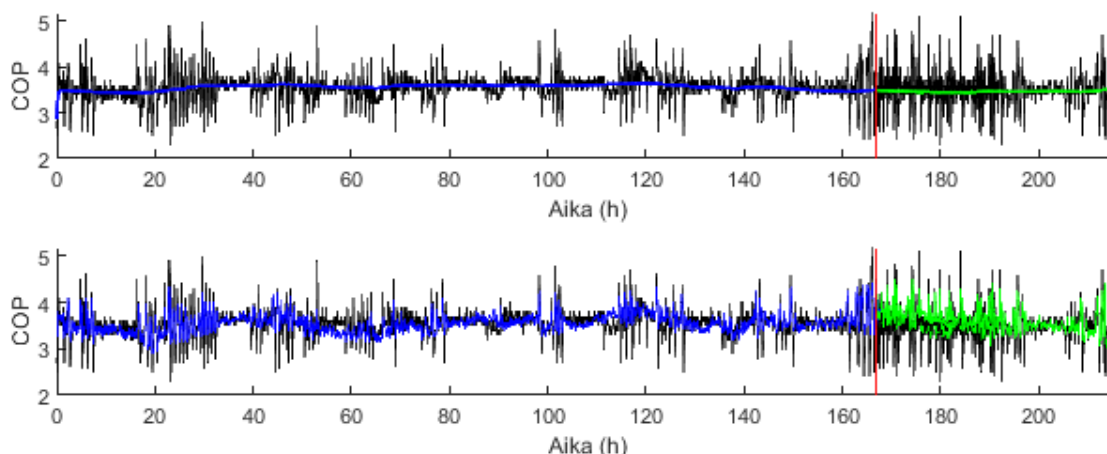


Kuva 9. Opetusjakson (sininen) ja ennusteen validointitulos (vihreä) dynaamisille regressiomalleille Q2 (ylhällä), Q3 (keskellä) ja Q4 (alhaalla) mittausaineiston 1 testidatalla. Musta viiva on lämpökertoimen simuloitu referenssiarvo.

5.3.2 Regressio- ja neuroverkkomallit: mittausaineisto

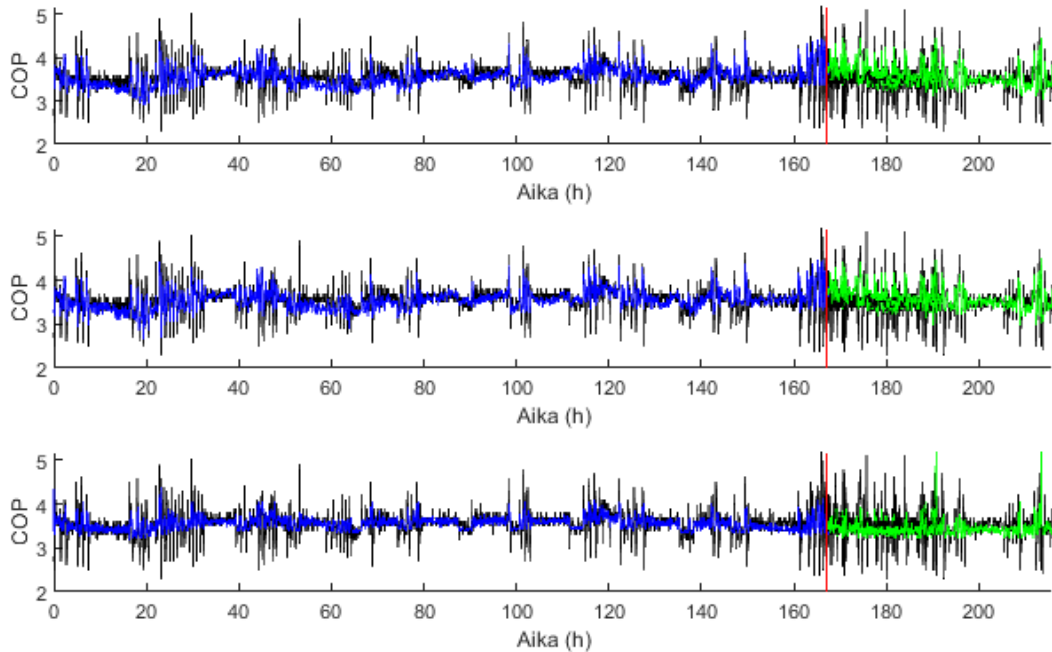
Regressiomalleille määritettiin opetusjakson pituudeksi 14 vuorokautta ja 7 vuorokautta. Keskimääräiset RMSE-arvot laskettiin ja niiden perusteella tultiin siihen johtopäätökseen, että 7 vuorokauden opetusjakso tuottaa parempia tuloksia.

Kuvassa 10 näkyy mittausaineistolla 2 identifioitujen lineaaristen mallien L1 ja L2 tulokset, jotka on saatu käyttämällä malliin ensimmäistä 7 vuorokauden opetusjaksoa. Mallista L1 näkee, että pelkällä ulkolämpötilalla on hyvin vähäinen merkitys poistoilmalämpöpumpun COP-arvoon, sillä niin opetusjoukko kuin testijoukkokaan eivät seuraa riittävällä tarkkuudella referenssiarvoa minuuttitasolla. Malli kuitenkin pystyy seuraamaan pidemmän aikavälin trendiä. Mallista L2 näkee selvän eron L1-malliin verrattuna, sillä malli kykenee ottamaan paremmin huomioon suurempia vaihteluita COP-arvossa. Paluuveden lämpötilalla on siten tärkeä merkitys lämpökertoimen ennustamisessa. L1:n RMSE ja MAPE olivat opetusjaksolle 0,33 ja 6,68 % sekä testidatalla 0,44 ja 9,36 %.



Kuva 10. Opetusjakson (sininen) ja ennusteen validointitulokset (vihreä) dynaamisille regressiomalleille L1 (ylempi) ja L2 (alempi) mittausaineistolla 2. Musta viiva on lämpökertoimentoteutunut arvo.

Kuva 11 esittää mitatulla datalla identifioituja neliöllisten mallien Q2, Q3 ja Q4 tuloksia, kun mallin toteutukseen on käytetty ensimmäistä datajoukkoa. Malli Q2 on samankaltainen L2-mallin kanssa, sillä mallien opetukseen on käytetty samoja tulomuuttujia. Q2-malli on kuitenkin hieman tarkempi, jonka näkee taulukon 5 tilastollisista mittareista. Ero on vain pieni, sillä esimerkiksi RMSE:n erotus on vain 0,01 yksikköä neliöllisen mallin hyväksi. Näyttäisi siis siltä, että neliöllisten termien lisääminen parantaa hieman mallin tarkkuutta. Malli Q3 on tarkasteltavista regressiomalleista selvästi tarkin. Sen RMSE opetus- ja testijaksolle oli vain 0,23 ja 0,32 ja MAPE 4,62 % ja 6,13 %. Syynä on varmasti, että se sisältää eniten muuttujia ja neliöllisiä termejä. Malli pystyy kuvaamaan kohtalaisen hyvin suuriakin COP-arvon vaihteluita. Malli Q4 on muihin tarkasteltuihin regressiomalleihin verrattuna heikoin. Koska Q2-malli (RMSE 0,32) on tarkempi kuin Q4-malli (RMSE 0,41), voidaan päätellä, että paluuveden lämpötilalla on suurempi merkitys kuin kompressorin hetkellisellä teholla poistoilmalämpöpumpun lämpökertoimen ennusteeseen.



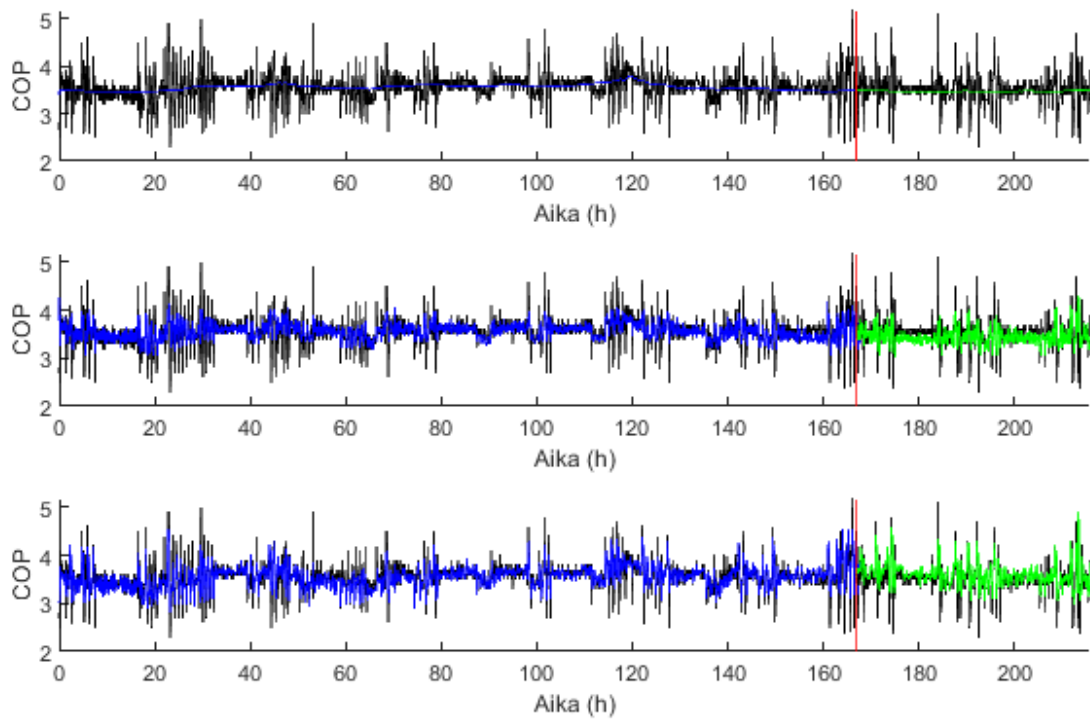
Kuva 11. Opetusjakson (sininen) ja ennusteen validointitulos (vihreä) neliöllisille dynaamisille regressiomalleille Q2 (ylhällä), Q3 (keskellä) ja Q4 (alhaalla) mittausaineistolla 2. Musta viiva on lämpökertoimen toteutunut arvo.

Neuroverkkomallien neuronien lukumäärä optimoitiin testaamalla neuronien eri lukumääriä ja minimoimalla testijoukkojen keskimääräinen MSE. Virhe vaihteli hieman eri testauskerroilla, joten pienin virhe voitiin saada eri neuronien määrällä kuin aiemmin. Kun neuronien määrä oli alle 10, virheissä ei ollut kuitenkaan merkittävän suurta eroa. Mallit eivät kunnolla konvergoituneet, jos neuroneita oli enemmän kuin 10 ja laskenta-aika kasvoi, mitä enemmän neuroneita oli. Työssä esitetyissä neuroverkkomallien tuloksissa on käytetty kahta neuronaa NARX1:lle ja NARX2:lle ja kolmea neuronaa NARX3:lle. Taulukosta 4 nähdään, että NARX1:n pienin keskimääräinen neliövirhe pysyi lähes samana (0,25), kun neuroneita oli yhdestä kuuteen. NARX2:n pienin neliövirhe saatiin kahdella neuronilla (0,22) ja NARX3:n tapauksessa kolmella neuronilla (0,14).

Taulukko 4. Kolmen mitatun datajoukon keskimääräinen neliövirhe (MSE) eri neuronien lukumäärällä ja NARX1, NARX2 ja NARX3 mallirakenteilla. Pienimmät virheen arvot on korostettu.

| Malli | Neuronit (lkm) | | | | | | | | | |
|-------|----------------|--------------|--------------|-------|--------------|--------------|-------|-------|------------------------|------------------------|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| NARX1 | 0,251 | 0,250 | 0,250 | 0,251 | 0,250 | 0,250 | 0,406 | 0,251 | ei konvergoi -du | ei konvergoi- du |
| NARX2 | 0,224 | 0,217 | 0,226 | 0,218 | 0,217 | 0,228 | 0,228 | 0,221 | 0,219 | 0,224 |
| NARX3 | 0,147 | 0,145 | 0,139 | 0,148 | 0,148 | 0,140 | 0,141 | 0,148 | 0,142 | 0,141 |

Kuvassa 12 näkyvät kolmen testatun neuroverkkomallin tulokset ensimmäiselle datajoukolle kolmesta. Kuvasta 12 voidaan havaita, että mallin tarkkuus paranee, kun muuttujien lukumäärä kasvaa. Ulkolämpötilalla ei ole kovin paljon vaikutusta minuuttitasolla. Se on tässä tapauksessa ymmärrettävää, sillä malleissa käytetty data on peräisin poistoilmalämpöpumpusta. Kun mukaan lisätään kompressorin hetkellinen teho, tarkkuus paranee, mutta NARX2 malli ei kykene ennustamaan yhtä tarkasti suurimpia piikkejä kuin NARX3.



Kuva 12. Opetusjakson (sininen) ja ennusteen validointitulokset (vihreä) neuroverkkomalleille NARX1 (ylhällä), NARX2 (keskellä) ja NARX3 (alhaalla) mittausaineistolla 2. Musta viiva on lämpökertoimen mitattu referenssiarvo.

Taulukkoon 5 on koottu mittausaineistolla 2 testattujen dynaamisten regressiomallien sekä neuroverkkomallien kolmen datajoukon keskimääräiset RMSE, MAPE ja korrelaatio sekä niiden keskihajonnat. Parhaiten suoriutuva regressiomalli saadaan monimutkaisimmalla Q3-mallilla, joka sisältää eniten termejä ja kaikki muuttujat. Epätarkin on mallirakenne L1, joka huomioi vain lineaarisen riippuvuuden ulkolämpötilaan. Neuroverkkomalleista paras malli oli NARX3, jonka RMSE ja MAPE olivat opetusjoukolle 0,23 ja 4,43 % ja testiosalle 0,32 ja 6,06 %. NARX3 oli vertailussa samaa luokkaa tai hieman parempi kuin regressiomalli Q3, jonka RMSE testidatalla oli 0,32.

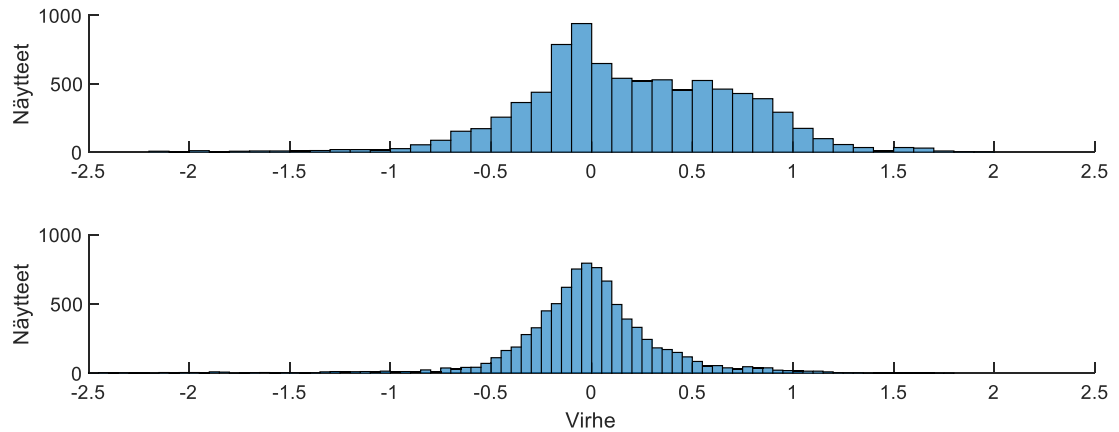
Taulukko 5. Keskimääräiset RMSE, MAPE ja korrelaatio sekä niiden keskihajonnat kolmella mitatulla datajoukolla dynaamisille regressiomalleille ja neuroverkoille.

| Malli | Opetusdata (14 vrk tai 7vrk*) | | | Testidata (2 vrk) | | |
|--------|-------------------------------|------------|------------|-------------------|------------|------------|
| | RMSE | MAPE (%) | r | RMSE | MAPE (%) | r |
| L1 | 0,412±0,08 | 8,150±1,52 | 0,341±0,06 | 0,473±0,17 | 9,785±3,15 | 0,273±0,04 |
| L2 | 0,325±0,07 | 6,206±1,05 | 0,675±0,00 | 0,378±0,14 | 7,243±2,03 | 0,654±0,05 |
| Q2 | 0,315±0,06 | 6,010±0,92 | 0,699±0,01 | 0,369±0,15 | 7,161±2,12 | 0,677±0,06 |
| Q3 | 0,299±0,06 | 5,546±0,78 | 0,733±0,01 | 0,351±0,14 | 6,564±1,88 | 0,702±0,05 |
| Q4 | 0,384±0,08 | 7,293±1,43 | 0,496±0,02 | 0,450±0,18 | 9,145±3,12 | 0,543±0,02 |
| L1* | 0,330±0,04 | 6,675±1,04 | 0,361±0,12 | 0,435±0,17 | 9,357±4,87 | 0,059±0,05 |
| L2* | 0,256±0,04 | 5,155±0,83 | 0,704±0,00 | 0,336±0,15 | 6,632±2,84 | 0,647±0,04 |
| Q2* | 0,251±0,04 | 5,049±0,74 | 0,719±0,00 | 0,324±0,14 | 6,306±2,62 | 0,651±0,03 |
| Q3* | 0,234±0,03 | 4,623±0,48 | 0,756±0,02 | 0,318±0,13 | 6,127±2,35 | 0,676±0,03 |
| Q4* | 0,287±0,03 | 5,548±0,55 | 0,585±0,08 | 0,414±0,20 | 8,372±4,96 | 0,337±0,14 |
| NARX1 | 0,411±0,08 | 8,079±1,56 | 0,365±0,03 | 0,472±0,17 | 9,747±3,09 | 0,222±0,09 |
| NARX2 | 0,367±0,08 | 7,055±1,39 | 0,550±0,02 | 0,440±0,19 | 8,900±3,41 | 0,491±0,16 |
| NARX3 | 0,277±0,06 | 5,320±0,86 | 0,755±0,01 | 0,345±0,15 | 6,457±2,08 | 0,705±0,06 |
| NARX1* | 0,328±0,04 | 6,585±0,81 | 0,378±0,12 | 0,430±0,18 | 9,268±4,22 | 0,062±0,06 |
| NARX2* | 0,277±0,03 | 5,331±0,42 | 0,625±0,07 | 0,400±0,20 | 8,037±3,87 | 0,401±0,16 |
| NARX3* | 0,225±0,03 | 4,426±0,45 | 0,780±0,01 | 0,316±0,13 | 6,061±2,08 | 0,685±0,02 |

5.4 Mallinnusvirheen analyysi

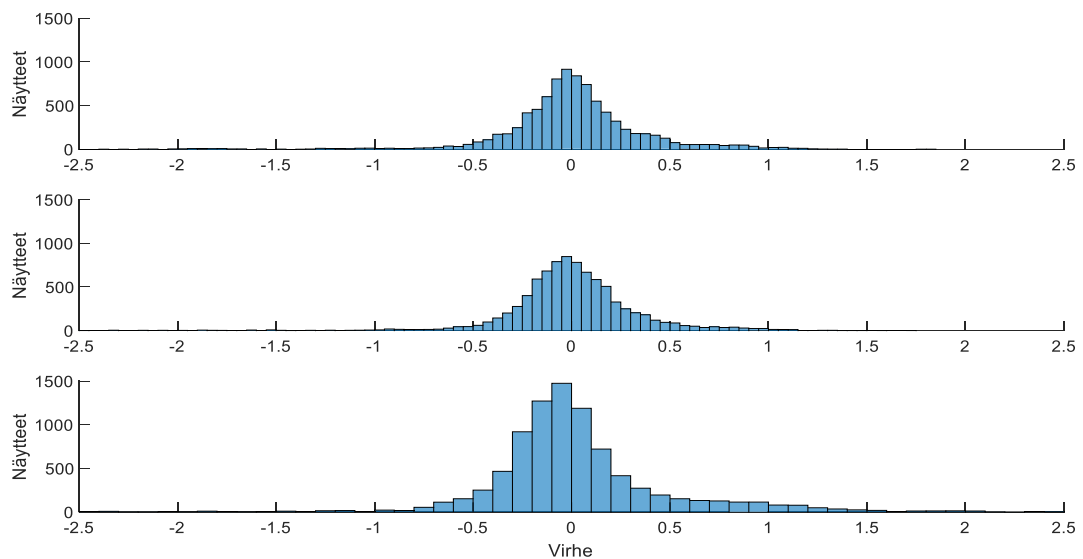
Mittausaineistolla 2 mallinnusvirheitä analysoitiin testidatan virheen jakauman avulla. Mallinnusvirheen histogrammi kuvaa lähtöarvon ennusteen ja referenssiarvon välisen erotuksen jakaumaa. Positiivinen virhe tarkoittaa, että lähtöarvo oli suurempi kuin referenssiarvo ja negatiivinen, että referenssiarvo oli suurempi kuin lähtöarvo. Kun histogrammi on symmetrinen, virhe jakautuu tasaisesti molemmiin puoliin nollaa.

Kuvassa 13 on dynaamisten regressiomallien L1 ja L2 mallinnusvirheen jakauma histogrammina. Mallin L2 virhejakauma on symmetrisempi ja virhejakauma on pienempi kuin L1:n, joten sen virheen jakauma on tasaisempi ja sisältää vähemmän virhettä.



Kuva 13. Dynaamisten lineaaristen regressiomallien L1 (ylhäällä) ja L2 (alhaalla) testidatojen mallinnusvirheen histogrammit mittausaineistolla 2.

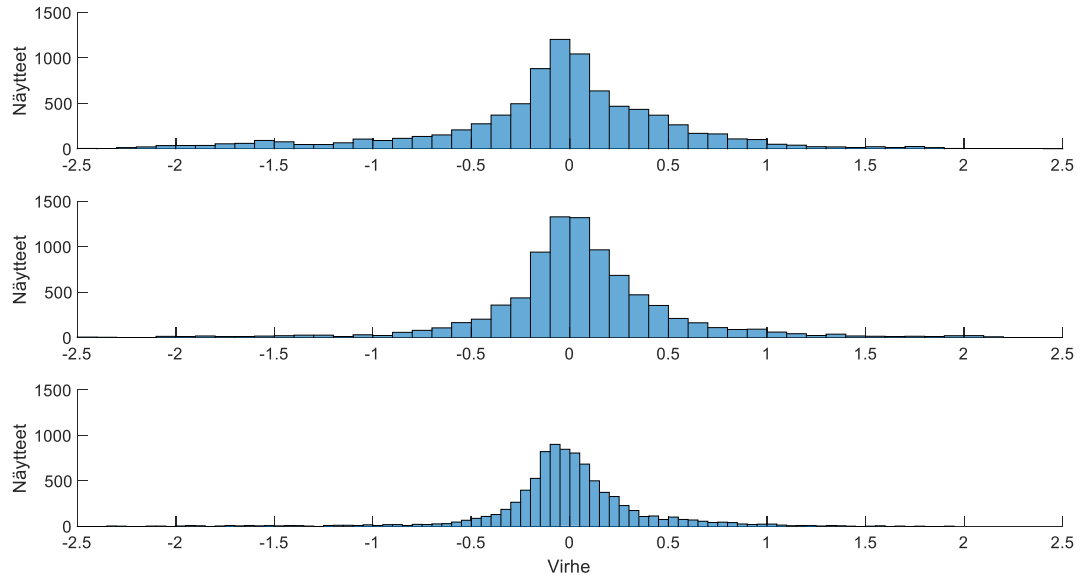
Kuvassa 14 on neliöllisten dynaamisten regressiomallien Q2, Q3, ja Q4 virhejakauman histogrammit mittausaineistolla 2. Virhejakauma on kaikissa lähes symmetrinen, mutta kuitenkin painottuu enemmän positiiviselle puolelle, joten lähtöarvot ylittävät todennäköisemmin referenssiarvon kuin toisinpäin. Kapein virhejakauma on mallirakenteella Q3.



Kuva 14. Dynaamisten neliöllisten regressiomallien Q2 (ylhäällä), Q3 (keskellä) ja Q4 (alhaalla) testijoukkojen mallinnusvirheen histogrammit mittausaineistolla 2.

Kuvassa 15 on puolestaan neuroverkkomallien NARX1, NARX2 ja NARX3 mallinnusvirheen jakauma. NARX1:llä on neuroverkoista suurin virhejakauma ja virhe

painottuu hieman enemmän negatiivisella puolella, joten lähtöarvot ovat keskimäärin pienempiä kuin referenssiarvot. NARX2 ja NARX3 jakautuvat suhteellisen symmetrisestä nollan molemmin puolin, joten virhettä on tasaisesti molemmin puolin. NARX3:n virhejakauma on kapein ja symmetrisin, joten se sisältää vähiten mallinnusvirhettä COP-arvon ennusteissa.



Kuva 15. Neuroverkkomallien NARX1 (ylhäällä), NARX2 (keskellä) ja NARX3 (alhaalla) testijoukkojen mallinnusvirheen histogrammit mittausaineistolla 2.

6 JOHTOPÄÄTÖKSET JA SUOSITUKSET

Mittausaineiston 1 ja mittausaineiston 2 tuloksia ei voida suoraan verrata keskenään, sillä niiden resoluutio ja laajuus olivat erilaisia. Lisäksi ensimmäinen oli peräisin ilma-vesilämpöpumpusta ja toinen poistoilmalämpöpumpusta.

Mittausaineistolla 2 testattu opetusjakson pituus vaikutti siten, että lyhyempi 7 vuorokauden opetusjakso osoittautui paremmaksi vaihtoehdoksi kuin 14 vuorokautta, sillä testijaksoille lasketut RMSE ja MAPE olivat pienempiä ja korrelaatiot suurempia lyhyemmällä opetusjaksolla. Tulos on looginen, sillä lähihistorialla on varmasti enemmän vaikutusta myös tulevaan lämpökertoimen arvoon.

Parhaimmaksi regressiomalliksi osoittautui neliöllinen kaikki muuttujat (ulkolämpötila, kompressorin hetkellinen teho ja latauspiirin paluuveden lämpötila) sisältävä Q3, jonka mittausaineistolla 2 saatu keskimääräinen RMSE ja keskihajonta oli testijaksolle $0,32 \pm 0,13$, MAPE oli puolestaan $6,13 \pm 2,35$ % sekä korrelaatio oli $0,68 \pm 0,03$, kun mallien opetuksessa käytettiin 7 vuorokauden opetusjaksoa.

Mallirakenne Q2 (RMSE $0,32 \pm 0,14$) suoriutui mittausaineistolla 2 suorituskymmittareiden mukaan paremmin kuin Q4 (RMSE $0,41 \pm 0,2$), joten sen perusteella latauspiirin paluuveden lämpötila vaikuttaa mallin tarkkuuteen enemmän kuin kompressorin hetkellinen teho. Yllättävää oli kuitenkin, että mittausaineistolla 1 saatu Q4 oli tarkempi (RMSE 0,09) kuin malli Q2 (RMSE 0,2). Tuntitason datalla saadut mallien virheet olivat pienempiä kuin minuuttitason. Koska mittausaineiston 2 data oli peräisin poistoilmalämpöpumpusta, ulkolämpötilalla oli vain vähän merkitystä.

Myös neuroverkkomallien kohdalla kaikki kolme muuttujaa sisältävä NARX3 osoittautui parhaaksi. Sen mittausaineistolla 2 saatu keskimääräinen RMSE oli testijaksolle 0,32, MAPE oli 6,06 % ja ennusteen korrelaatio oli 0,69. Näiden suorituskymmittareiden mukaan neuroverkkomallit olivat siis hieman tarkempia kuin dynaamiset regressiomallit. Neuroverkot sisältävät takaisinkytkennän, joten ehkä tästä syystä niiden ennustetarkkuus on parempi. Toisaalta neuroverkkojen mallirakenne on monimutkaisempi ja ne vaativat enemmän laskentakapasiteettia, jolloin laskenta tapahtuu hitaammin ja malli ei välttämättä konvergoitu kunnolla.

Mallien identifiointiin käytetyn datan ajankohdalla oli vaikutusta mallin tarkkuuteen. Helmikuun datalla koulutettujen mallien virheet olivat pienempiä kuin maaliskuun. Tämä selittynee sillä, että maaliskuun datassa oli huomattavasti suurempaa vaihtelua eri muuttujissa etenkin paluuveden lämpötilan osalta. Vaihtelua oli tosin maaliskuussa niin runsaasti, että se voi johtua osittain mittausvirheestä. Myös ulkolämpötila oli maaliskuussa korkeampi kuin helmikuussa. Siksi jatkon kannalta olisi hyvä testata, kuinka hyvin mallit ovat yleistettävissä eri vuodenaikoihin. Tähän tarvitaan laajempi datajoukko. Tulosten perusteella voidaan arvioida, että testatut mallit ovat hyvin sovellettavissa myös muille lämpöpumpputyypeille. Ainakin regressiomallit toimivat hyvin sekä mittausaineistolla 1, joka oli ilma-vesilämpöpumpusta, että poistoilmalämpöpumpun mittausaineistolla 2. Laajemman yleistettävyyden varmistamiseksi tarvitaan lisää tutkimusta.

Tuntitason datalla (mittausaineisto 1) saatu parhaan mallin Q3 testijakson MAPE oli 0,57 % ja korrelaatio 0,96. Mittausaineistolla 2 parhaiden mallien (NARX3 ja Q3) MAPE olivat 6,06 % ja 6,13 % sekä korrelaatio 0,69 ja 0,68. Etenkin mittausaineistolla 1 saadut tulokset pärjäävät vertailussa kappaleessa 4 esitettyjen aiempien tutkimusten arvoihin. Esimerkiksi Lu ym. (2019) RF-menetelmällä saatu MAPE oli 1,15–1,93 % ja korrelaatio 0,94–0,98. Toisaalta korrelaatiot jäivät vertailussa pienemmäksi, sillä useassa tutkimuksessa (Esen ym. 2008a; Sun ym. 2015; Zhang ym. 2013; Fannou ym. 2014; Shin ym. 2019; Zou & Xie ym. 2017; Simon ym. 2016) korrelaatio oli erittäin korkea, noin 0,98–0,99. Suoraan vertailuun tulee kuitenkin suhtautua varauksella, sillä käytetyn datan laadussa, resoluutiossa ja laajuudessa sekä mallirakenteiden monimutkaisuudessa on varmasti eroja. Mittausaineistolla 2 saatuihin arvoihin ei ole kovin mielekäästä verrata, sillä käsittääkseni yksikään muu tutkimus ei ollut toteutettu minuuttitason datalla, kuten tässä työssä. Lisäksi aiemmat tutkimukset liittyivät käytännössä kaikki maalämpöpumppuihin tai ilma-vesilämpöpumppuihin, kun taas mittausaineisto 2 oli poistoilmalämpöpumpusta.

Lämpökerroin on yksi tärkeimpiä lämpöpumpun toimintaa ja tehokkuutta kuvaava mittari. Tulosten perusteella työssä käytetyillä malleilla voidaan ennustaa riittävällä tarkkuudella ilma-vesilämpöpumpun ja poistoilmalämpöpumpun lämpökerrointa jo kolmella muuttujalla (ulkolämpötila, paluuveden lämpötila ja kompressorin hetkellinen teho).

Jatkossa malleja voidaan edelleen parantaa esimerkiksi lisäämällä muuttujien määrää (esimerkiksi keruupiirin lämpötila) tai kokeilemalla eri muuttujakombinaatioita tai muodostamalla matemaattisia operaatioita hyödyntäen uusia muuttujia. Myös erilaisia mallirakenteita voidaan testata, sillä niitä on lukuisia. Neuroverkkomallien parametrien viritystä voidaan hienosäätää ja kokeilla erilaisia opetusalgoritmeja.

Jatkossa voidaan myös esimerkiksi miettiä, voisiko ennustemalleja hyödyntää lämpöpumppujen operoinnin energiatehokkaaseen säätöön tai mahdollisesti optimoida lämpöpumpun operointia kysyntäjouaston tarpeen mukaan.

7 YHTEENVETO

Tässä työssä tutkittiin datapohjaisia mallinnusmenetelmiä lämpöpumpun lämpökertoimen ennustamisessa. Kirjallisuuskatsauksen perusteella mallinnus aiemmissa vastaavissa tutkimuksissa painottui vahvasti maalämpöpumpppuihin ja niissä käytetyimmät menetelmät olivat enimmäkseen neuroverkkopohjaisia lämpöpumpppuprosessin epälineaarisen luonteen vuoksi. Yleisimmin lämpökertoimen ennustamisessa tulomuuttujina käytettiin ulkolämpötilaa, paluuveden lämpötilaa ja tehoa, joten ne valittiin tulomuuttujiksi myös työssä kehitettyihin malleihin.

Kokeellisen osan malleja testattiin kahdella erityyppisellä mittausaineistolla. Työssä identifioitiin ja validoitiin viisi erilaista dynaamista regressiomallia ja kolme epälineaarista eksogeenista autoregressiivistä neuroverkkomallia. Tulosten perusteella parhaiten ennustavat mallirakenteet olivat epälineaarinen autoregressiivinen neuroverkkomalli ja dynaaminen neliöllinen regressiomalli. Molemmat sisälsivät kaikki kolme käytössä ollutta tulomuuttujaa. Mittausaineistolla 2 saadun parhaan neuroverkkomallin keskimääräiset tunnusluvut olivat testidatalla 0,32, 6,06 % ja 0,69 ja regressiomallin vastaavasti 0,32, 6,13 % ja 0,68. Mittausaineistolla 1 parhaan regressiomallin keskimääräinen keskineliövirheen neliöjuuri oli 0,04, keskimääräinen absoluuttinen suhteellinen virhe 0,57 % ja korrelaatio 0,96.

Tarkimmat mallit saatiin käyttämällä kaikkia käytössä ollutta kolmea tulomuuttujaa. Samalla kuitenkin mallien monimutkaisuus kasvoi. Regressiomallien suorituskkyä paransi neliöllisten termien lisääminen. Jatkon kannalta ehkä oleellisinta on tutkia, millaisissa käytännön sovelluksissa mallia voitaisiin hyödyntää. Kun lämpöpumpun toimintaa pystytään ennustamaan jo etukäteen, voidaan mahdollisesti kehittää mallia hyödyntävä säätöstrategia energiatehokkuuden parantamiseksi tai optimoida lämpöpumpun toimintaa esimerkiksi kysyntäjouston suhteen.

Kaiken kaikkiaan sekä mallien opetusjakso että testijakso seurasivat riittävällä tarkkuudella referenssiarvoa, joten työssä toteutettuja datapohjaisia lämpökertoimen ennustusmalleja on mahdollista käyttää lämpöpumpun toiminnan ennustamiseen.

LÄHDELUETTELO

- Afram, A. & Janabi-Sharifi, F., 2015. Black-box modeling of residential HVAC system and comparison of gray-box and black-box modeling methods. *Energy and Buildings*, 94, 121–149. <https://doi.org/10.1016/j.enbuild.2015.02.045>
- Amasyali, K. & El-Gohary, N. M., 2018. A review of data-driven building energy consumption prediction studies. *Renewable and Sustainable Energy Reviews*, 81, 1192–1205. <https://doi.org/10.1016/j.rser.2017.04.095>
- Biau, G., ja Scornet, E., 2016. A random forest guided tour. *TEST*, 25 (2), 197–227. <https://doi.org/10.1007/s11749-016-0481-7>
- Bishop, C. M., 2006. *Pattern recognition and machine learning*. New York: Springer, 738 s. ISBN 9780387455280
- Bourdeau, M., Zhai, X., Nefzaoui, E., Guo, X. & Chatellier, P., 2019. Modeling and forecasting building energy consumption: A review of data-driven techniques. *Sustainable Cities and Society*, 48, 101533. <https://doi.org/10.1016/j.scs.2019.101533>
- Cupeiro Figueroa, I., Picard, D. & Helsen, L., 2020. Short-term modeling of hybrid geothermal systems for Model Predictive Control. *Energy and Buildings*, 215, 109884. <https://doi.org/10.1016/j.enbuild.2020.109884>
- Esen, H., Inalli, M., Sengur, A. & Esen, M., 2008a. Forecasting of a ground-coupled heat pump performance using neural networks with statistical data weighting pre-processing. *International Journal of Thermal Sciences*, 47, 431–441. <https://doi.org/10.1016/j.ijthermalsci.2007.03.004>
- Esen, H., Inalli, M., Sengur, A. & Esen, M., 2008b. Artificial neural networks and adaptive neuro-fuzzy assessments for ground-coupled heat pump system. *Energy and Buildings*, 40, 1074–1083. <https://doi.org/10.1016/j.enbuild.2007.10.002>

Esen, H., Inalli, M., Sengur, A. & Esen, M., 2008c. Predicting performance of a ground-source heat pump system using fuzzy weighted pre-processing-based ANFIS. *Building and Environment*, 43, 2178–2187. <https://doi.org/10.1016/j.buildenv.2008.01.002>

Esen, H., Inalli, M., Sengur, A. & Esen, M., 2008d. Modeling a ground-coupled heat pump system by a support vector machine. *Renewable Energy*, 33, 1814–1823. <https://doi.org/10.1016/j.renene.2007.09.025>

Esen, H. & Inalli, M., 2010. ANN and ANFIS models for performance evaluation of a vertical ground source heat pump system. *Expert Systems with Applications*, 37, 8134–8147. <https://doi.org/10.1016/j.eswa.2010.05.074>

Fannou, J., Rousseau, C., Lamarche, L. & Kajl, S., 2014. Modeling of a direct expansion geothermal heat pump using artificial neural networks. *Energy and Buildings*, 81, 381–390. <https://doi.org/10.1016/j.enbuild.2014.06.040>

Ferrarini, L., Rastegarpour, S., Caseri, L., Watanabe, T. & Choi, Y., 2020. Predictive control-oriented models of a domestic air-to-water heat pump under variable conditions. *IEEE Robotics and Automation Letters*, 4, 5363–5369. <https://doi.org/10.1109/lra.2020.3007474>

Jesper, M., Schlosser, F., Pag, F., Walmsley, T.G., Schmitt, B., Vajen, K., 2021. Large-scale heat pumps: Uptake and performance modelling of market-available devices. *Renewable and Sustainable Energy Reviews* 137, 110646. <https://doi.org/10.1016/j.rser.2020.110646>

Lu, S., Li, Q., Bai, L. & Wang, R., 2019. Performance predictions of ground source heat pump system based on random forest and back propagation neural network models. *Energy Conversion and Management*, 197, 111864. <https://doi.org/10.1016/j.enconman.2019.111864>

Maa- ja metsätalousministeriö, 2021. EU:n energia- ja ilmastopolitiikka [verkkodokumentti]. Helsinki. Saatavissa: <https://mmm.fi/luonto-ja-ilmasto/energia-ja-ilmastopolitiikka/eu-energia-ja-ilmastopolitiikka> [viitattu 9.4.2021]

Markovic, R., Lorz, C., Frisch, J. & van Treeck, C., 2017. Application of Support Vector Machines for Predicting the Performance of Air-Source Domestic Hot Water Heat Pump Systems. Building Simulation Proceedings of the 15th IBPSA Conference, San Francisco, CA, USA. 10 s. <https://doi.org/10.26868/25222708.2017.156>

MathWorks, 2020. Introduction to Feature Selection [verkkodokumentti]. Saatavissa: <https://se.mathworks.com/help/stats/feature-selection.html> [viitattu 15.2.2021]

MathWorks, 2021. Introducing Machine Learning [verkkodokumentti]. Saatavissa: <https://www.mathworks.com/content/dam/mathworks/ebook/gated/machine-Learning-ebook.pdf> [viitattu 16.2.2021]

Mohanraj, M., Jayaraj, S. & Muraleedharan, C., 2012. Applications of artificial neural networks for refrigeration, air-conditioning and heat pump systems—A review. Renewable and Sustainable Energy Reviews, 16, 1340–1358. <https://doi.org/10.1016/j.rser.2011.10.015>

Motiva, 2012a. Lämpöä omasta maasta – Maalämpöpumput [verkkodokumentti]. Saatavissa: https://www.motiva.fi/files/7965/Lampoa_omasta_maasta_Maalampopumput.pdf [viitattu 3.2.2021]. 16 s.

Motiva, 2012b. Lämpöä ilmassa – Ilmalämpöpumput [verkkodokumentti]. Saatavissa: <http://www.motiva.fi/files/175/Ilmalampopumput.pdf> [viitattu 5.2.2021]. 16 s.

Motiva, 2018. Lämpöpumppujen hankintaopas – kunnat ja taloyhtiöt [verkkodokumentti]. Saatavissa: https://www.motiva.fi/files/14752/Lampopumppujen_hankintaopas_kunnat_ja_taloyhti_ot.pdf [viitattu 8.2.2021]. Helsinki. 36 s.

Motiva, 2020. Lämpöpumput [verkkodokumentti]. Saatavissa: https://www.motiva.fi/ratkaisut/uusiutuva_energia/lampopumput [viitattu 13.2.2021]

Rastegarpour, S., Caseri, L., Ferrarini, L. & Gehrke, O., 2019. Experimental Validation of the Control-Oriented Model of Heat Pumps for MPC Applications. 2019 IEEE 15th

International Conference on Automation Science and Engineering (CASE). <https://doi.org/10.1109/coase.2019.8843211>

Rastegarpour, S., Gros, S. & Ferrarini, L., 2020. MPC approaches for modulating air-to-water heat pumps in radiant-floor buildings. *Control Engineering Practice*, 95, 104209. <https://doi.org/10.1016/j.conengprac.2019.104209>

Rastegarpour, S., Scattolini, R. & Ferrarini, L., 2021. Performance improvement of an air-to-water heat pump through linear time-varying MPC with adaptive COP predictor. *Journal of Process Control*, 99, 69–78. <https://doi.org/10.1016/j.jprocont.2021.01.006>

Reynolds, J., Ahmad, M. W. & Rezgui, Y., 2018. Holistic modelling techniques for the operational optimisation of multi-vector energy systems. *Energy and Buildings*, 169, 397–416. <https://doi.org/10.1016/j.enbuild.2018.03.065>

ScienceDirect, 2021. Topics: Feature Extraction [verkkodokumentti]. Saatavissa: <https://www.sciencedirect.com/topics/engineering/feature-extraction> [viitattu 9.9.2021]

Shin, J.-H., Kim, Y.-I. & Cho, Y.-H., 2019. Development of Operating Method of Multi-Geothermal Heat Pump Systems Using Variable Water Flow Rate Control and a COP Prediction Model Based on ANN. *Energies*, 12, 3894. <https://doi.org/10.3390/en12203894>

Simon, F., Ordoñez, J., Reddy, T. A., Girard, A. & Muneer, T., 2016. Developing multiple regression models from the manufacturer's ground-source heat pump catalogue data. *Renewable Energy*, 95, 413–421. <https://doi.org/10.1016/j.renene.2016.04.045>

Srinivasan P., 2019. Interpretable Machine Learning: An attempt to demystify the black-box [verkkodokumentti]. Saatavissa: <https://medium.com/walmartglobaltech/accuracy-vs-interpretability-paradox-382803f6a99d> [viitattu 13.3.2021]

Sun, W., Hu, P., Lei, F., Zhu, N. & Jiang, Z., 2015. Case study of performance evaluation of ground source heat pump system based on ANN and ANFIS models. *Applied Thermal Engineering*, 87, 586–594. <https://doi.org/10.1016/j.applthermaleng.2015.04.082>

Suomen lämpöpumppuyhdistys ry, 2021. Lämpöpumput [verkkodokumentti]. Saatavissa: <https://www.sulpu.fi/lampopumput> [viitattu 5.2.2021]

Thermia, 2021. Maalämmön toimintaperiaate [verkkodokumentti]. Saatavissa: <https://www.thermia.fi/maalampo/maalampo1/miten-maalampo-toimii/> [viitattu 15.8.2021]

Yan, L., Hu, P., Li, C., Yao, Y., Xing, L., Lei, F. & Zhu, N., 2016. The performance prediction of ground source heat pump system based on monitoring data and data mining technology. *Energy and Buildings*, 127, 1085–1095. <https://doi.org/10.1016/j.enbuild.2016.06.055>

Ympäristöministeriö, 2012. Lämpöpumppujen energialaskentaopas [verkkodokumentti]. Saatavissa: <https://docplayer.fi/683121-Lampopumppujen-energiaskentaopas.html> [viitattu 9.2.2021]

Verhelst, C., Logist, F., Van Impe, J. & Helsen, L., 2012. Study of the optimal control problem formulation for modulating air-to-water heat pumps connected to a residential floor heating system. *Energy and Buildings*, 45, 43–53. <https://doi.org/10.1016/j.enbuild.2011.10.015>

VTT Expert Services Oy, 2018. Ilmalämpöpumpun Mitsubishi MSZ-LN25VGW + MUZ-LN25VGHZ toimintakoe matalissa ulkoilman lämpötiloissa ja sulatusjaksot sisältävä lämpökerroin. Testausseoloste. [verkkodokumentti]. Saatavissa: https://www.scanoffice.fi/wp-content/uploads/sites/11/2018/01/VTT_ln25.pdf [viitattu 17.8.2021]

Wang, Y., Li, W., Zhang, Z., Shi, J. & Chen, J., 2019. Performance evaluation and prediction for electric vehicle heat pump using machine learning method. *Applied Thermal Engineering*, 159, 113901. <https://doi.org/10.1016/j.applthermaleng.2019.113901>

Zhang, Y, Wang, G; Han, G., 2013. GCHP system optimal predictive control based on RBFNN and APSO algorithm. *Proceedings of the 32nd Chinese Control Conference*. IEEE, 2402-2406. ISBN 978-9-8815-6383-5

Zou, S. & Xie, X., 2017. Simplified model for coefficient of performance calculation of surface water source heat pump. *Applied Thermal Engineering*, 112, 201–207.
<https://doi.org/10.1016/j.applthermaleng.2016.10.081>